SCIENTIFIC AMERICAN

JANUARY 1995

$3.95
U.K. £2.75

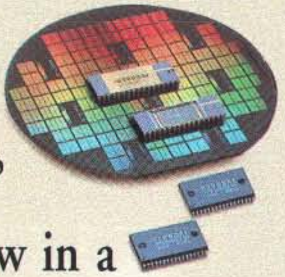
Warning: digital documents in danger.

Living well past age 100.

Laughing cannibals and mad cows.

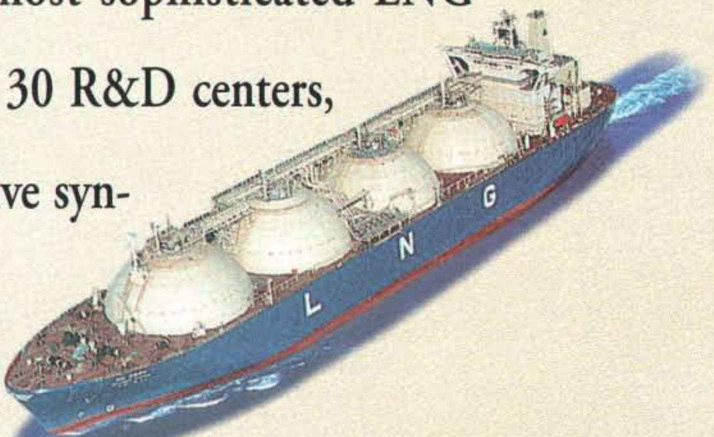
Supercontinent *of 750 million years ago*
is pieced together for the first time.

From Chips To Ships



Hyundai Business Group, with over US$58 billion in sales, continues to grow in a wide spectrum of business areas. From next generation 256M DRAM chips to satellite communications, creating tomorrow's global information superhighway. From a full

AUTOMOBILES, ELECTRONICS, SHIPBUILDING, ENGINEERING AND CONSTRUCTION, MACHINERY AND EQUIPMENT,PETROCHEMICA line of passenger cars to all types of commercial vehicles. From machine tools to the magnetic levitation train, the ideal mass transit system of the future. From turn-key engineering and construction projects to petrochemicals with advanced new material. From super tankers to some of the most sophisticated LNG carriers. Hyundai, with more than 30 R&D centers, working together, creating innovative synergies, innovative products.



HYUNDAI
Building A Better World Through Innovative Technology

RADING AND TRANSPORTATION. Find out how Hyundai can help you : Hyundai Corporation K.P.O. Box 92, Seoul, Korea. Tel +82-2-746-1873, Fax +82-2-741-2341.

SCIENTIFIC AMERICAN

January 1995 Volume 272 Number 1

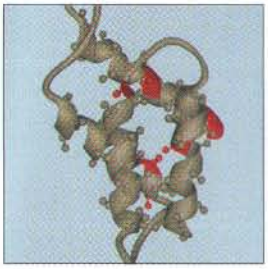
24

Ensuring the Longevity of Digital Documents

Jeff Rothenberg

Remember computer punch cards? How about data tapes? If you have entrusted valuable information to floppy disks, CD-ROMs or other digital media in the hope of preserving it for the ages, be forewarned. Changes in hardware and software technology can make digital documents unreadable. Prudent steps taken now, however, can guarantee that today's records will still be accessible tomorrow.

30



The Prion Diseases

Stanley B. Prusiner

Laughing cannibals and mad cows have something in common: both are often infected with pathogens called prions that cause neurodegenerative diseases in humans and animals. Unlike viruses or bacteria, prions do not contain genetic material. They are proteins that reshape a host's proteins into copies of themselves. The author of the prion concept explains how an outlandish idea came to be accepted.

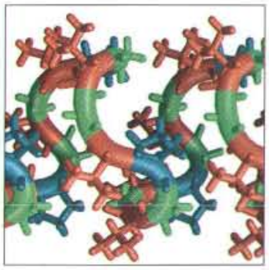
38

Earth before Pangea

Ian W. D. Dalziel

Today Antarctica is a frozen waste and California the Sunshine State, but 750 million years ago they appear to have been adjacent real estate. Long before the supercontinent Pangea coalesced 250 million years ago, plate tectonic forces were reshuffling landmasses and creating environments that nurtured primitive forms of life. Now geologists search for clues to the early wanderings of the continents.

44



Elastic Biomolecular Machines

Dan W. Urry

Taking a cue from the proteins in living things, chemists have begun to construct polymer molecules that expand or contract in response to changes in temperature, light or acidity. The thermodynamic qualities of these odd materials may seem perplexing. Yet they could be the key to building artificial muscles, new types of drug-delivery systems or more comfortable garments.

50

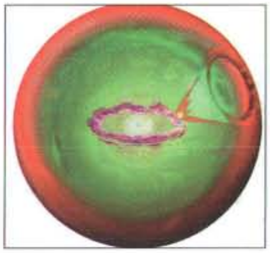
The Oldest Old

Thomas T. Perls

Many people regard advancing age as an inevitable descent into worsening health. A survey of persons who are more than 95 years old, however, finds that their physical condition is often better than that of others 20 years their junior. The longevity secrets locked inside these centenarians' genes and behavior may point the way to a more pleasurable and active old age for the rest of us.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 1994 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. Second-class postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian GST No. R 127387652. Subscription rates: one year $36 (outside U.S. and possessions add $11 per year for postage). Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send E-mail to SCAinquiry@aol.com.

56



The Birth and Death of Nova V1974 Cygni

Sumner Starrfield and Steven N. Shore

When this nova flared into existence in 1992, it was the brightest that astronomers had observed in 17 years. Then it unexpectedly faded, its nuclear fuels exhausted ahead of schedule. As astrophysicists pore over records of its brief life, V1974 Cygni is confirming some ideas about how stars explode but shattering others.

62

Egil's Bones

Jesse L. Byock

Old Norse tales commemorate the legendary accomplishments of the Viking hero Egil. They also note his appearance: his heavy features, his physical handicaps, his skull so thick it could withstand the blow of an ax. Such details may not be artistic embellishments. Egil may have had the skeletal condition called Paget's disease.

68

TRENDS IN PREVENTIVE MEDICINE

Better Than a Cure

Tim Beardsley, staff writer

Vaccines save millions of lives every year at very low cost, and the scientific prospects for creating more and better vaccines have never been brighter. But developing and distributing those drugs are still formidable jobs—ones that demand concerted effort by both industry and government. Can the World Health Organization break the logjam that has paralyzed vaccine progress in the past?

DEPARTMENTS

8

Science and the Citizen

Race and IQ.... Proteins that wire the brain.... Ozone lost between the poles.... Galactic magnetism.... Shoemaker-Levy's ongoing impact.... Mole rats: less naked but just as social.

The Analytical Economist
Derivatives and doom.

Technology and Business
Bellcore on the block?... Ownership of electronic art .... When 3-D pictures wore khaki.... Is there a doctor in the database?

Profile
Walter H. Munk, an oceanographer uneasily at odds with whale lovers.

5

Letters to the Editors
Orthopedics, 6,000 B.C.... The cause of the software crunch.

6

50 and 100 Years Ago
1945: Gyro navigation.
1895: The obsolete horse.

76

Mathematical Recreations
The dynamics of plant growth and the Fibonacci series.

80

Book Reviews
Architecture adapts.... Autofabrication.... Human genetics.

84

Essay: *John Timpane*
The irrational belief inside scientific conviction.

THE COVER painting portrays the earth 750 million years ago, when the major landmasses were fused into the supercontinent Rodinia. Except for parts of Africa and South America, there were no oceans between the continents. According to convention, they are shown to orient the viewer, as are the Great Lakes, Hudson Bay and Baffin Bay. North America eventually traveled around the other continental shields and rejoined them to form Pangea (see "Earth before Pangea," by Ian W. D. Dalziel, page 38). Painting by Tomo Narashima.

150 YEARS OF INNOVATION

Founded in 1845, as a weekly journal subtitled "The Advocate of Industry and Enterprise and Journal of Mechanical and Other Improvement," *Scientific American* is the oldest continuously published magazine in America.

From the beginning, *Scientific American* covered the new and important, most often in the words of those who actually did the work. Among the many people who have written for the magazine are Albert Einstein, Francis Crick and dozens of other Nobel laureates.

As the magazine enters its 150th year of innovation, readers and advertisers will be invited to participate in programs celebrating the sesquicentennial. Look for information in this space in subsequent issues, or check into *Scientific American*'s area on America Online.

SCIENTIFIC AMERICAN®
Established 1845

EDITOR IN CHIEF: John Rennie

BOARD OF EDITORS: Michelle Press, *Managing Editor;* Marguerite Holloway, *News Editor;* Timothy M. Beardsley; W. Wayt Gibbs; John Horgan, *Senior Writer;* Kristin Leutwyler; Philip Morrison, *Book Editor;* Madhusree Mukerjee; Sasha Nemecek; Corey S. Powell; Ricki L. Rusting; Gary Stix; Paul Wallich; Philip M. Yam

ART: Joan Starwood, *Art Director;* Edward Bell, *Art Director, Graphics Systems;* Jessie Nathans, *Associate Art Director;* Johnny Johnson, *Assistant Art Director, Graphics Systems;* Nisa Geller, *Photography Editor;* Lisa Burnett, *Production Editor*

COPY: Maria-Christina Keller, *Copy Chief;* Nancy L. Freireich; Molly K. Frances; Daniel C. Schlenoff

PRODUCTION: Richard Sasso, *Vice President, Production;* William Sherman, *Production Manager; Managers:* Carol Albert, *Print Production;* Janet Cermak, *Makeup & Quality Control;* Tanya DeSilva, *Prepress;* Silvia Di Placido, *Special Projects;* Carol Hansen, *Composition;* Madelyn Keyes, *Systems;* Ad Traffic: Carl Cherebin; Carey Ballard; Kelly Ann Mercado

CIRCULATION: Lorraine Leib Terlecki, *Associate Publisher/Circulation Director;* Katherine Robold, *Circulation Manager;* Joanne Guralnick, *Circulation Promotion Manager;* Rosa Davis, *Fulfillment Manager*

ADVERTISING: Kate Dobson, *Associate Publisher/Advertising Director.* OFFICES: NEW YORK: Meryle Lowenthal, *New York Advertising Manager;* Randy James, Rick Johnson, Elizabeth Ryan, Timothy Whiting. CHICAGO: 333 N. Michigan Ave., Suite 912, Chicago, IL 60601; Patrick Bachler, *Advertising Manager.* DETROIT: 3000 Town Center, Suite 1435, Southfield, MI 48075; Edward A. Bartley, *Detroit Manager.* WEST COAST: 1554 S. Sepulveda Blvd., Suite 212, Los Angeles, CA 90025; Lisa K. Carden, *Advertising Manager;* Tonia Wendt. 235 Montgomery St., Suite 724, San Francisco, CA 94104; Debra Silver. CANADA: Fenn Company, Inc. DALLAS: Griffith Group

MARKETING SERVICES: Laura Salant, *Marketing Director;* Diane Schube, *Promotion Manager;* Susan Spirakis, *Research Manager;* Nancy Mongelli, *Assistant Marketing Manager;* Ethel D. Little, *Advertising Coordinator*

INTERNATIONAL: EUROPE: Roy Edwards, *International Advertising Manager,* London; Vivienne Davidson, Linda Kaufman, Intermedia Ltd., Paris; Karin Ohff, Groupe Expansion, Frankfurt; Barth David Schwartz, *Director, Special Projects,* Amsterdam. SEOUL: Biscom, Inc. TOKYO: Nikkei International Ltd.; TAIPEI: Jennifer Wu, JR International Ltd.

ADMINISTRATION: John J. Moeling, Jr., *Publisher;* Marie M. Beaumonte, *General Manager*

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue, New York, NY 10017-1111
(212) 754-0550

CHAIRMAN AND CHIEF EXECUTIVE OFFICER: John J. Hanley

CO-CHAIRMAN: Dr. Pierre Gerckens

CORPORATE OFFICERS: John J. Moeling, Jr., *President;* R. Vincent Barger, *Chief Financial Officer;* Robert L. Biewen, *Vice President*

DIRECTOR, ELECTRONIC PUBLISHING: Martin Paul

PRINTED IN U.S.A.

LETTERS TO THE EDITORS

Picking at Bones

In "The Eloquent Bones of Abu Hureyra" [Scientific American, August], Theya Molleson spends an entire article addressing the amount of stress put on the bones of the people who lived in this agricultural village only to end by stating that "there was a constant progress toward a better life...." Implicit in this statement lies the ethnocentric assumption that the sedentary agricultural lifestyle is "better" than the nomadic hunter-gatherer lifestyle. It is the assumption that agricultural/industrial society is better than all others that has put us on our current path of environmental degradation. How many cultures and how many species must be lost before we realize that what we have is not better but simply different?

James Snell
Nashville, Tenn.

The suggestion made by Molleson that bone deformities resulted from grinding grain makes little sense to an orthopedic surgeon. No amount of grinding from an all-fours position would produce a marked hyperextended position of the metatarsal-phalangeal joint of the big toe unless the flexor profundus tendon were ruptured, most likely from a puncture or lacerating injury. Furthermore, arthritis of the metatarsal-phalangeal joint, known as hallux rigidus, can be traced to trauma but frequently is classified as idiopathic. The arthritic changes in the lumbar spine and the knee joint illustrated in the article are indistinguishable from those seen in many older persons in our society today.

Douglas B. Mains
Orthopaedic Associates of DuPage, Ltd.
Carol Stream, Ill.

Molleson replies:

Examination of the damaged metatarsal-phalangeal joint surface indicates that there was continued movement at the joint after the cartilage was damaged, which is consistent with repetitive minor trauma. Hallux rigidus following trauma was the differential diagnosis of a consulting radiologist. I interpret these injuries as having been sustained when the woman doing the grinding overshot the end of the quern. The toes would then be hyperextended beyond the normal range of movement, with the full driving force of the body behind them. Uninjured first metatarsals have an extension of the articular area of the head, that is, a kneeling facet.

Mains is correct that arthritic changes alone cannot tell a clear story. It is the association of spinal, knee and foot injuries in several individuals, where there is little other pathology, that suggests that they are consequent on a specific type of activity.

Software's Hard Questions

In "Software's Chronic Crisis" [Scientific American, September], W. Wayt Gibbs theorizes disaster for software development without the introduction of scientific methods and mathematical rigor. One interesting point is the mention of the standard engineering handbook approach, successful in many branches of engineering. Unfortunately, that approach works only with routine designs in well-established fields, by well-trained people making explicit and limiting assumptions. How many of these factors exist in real-world software projects? It is miraculous that large software systems can be built at all. In the future, people will look back in amazement at the large software projects done in the "prehistoric age" of the computer.

Robert G. Brown
Orange Park, Fla.

The "software crisis" we are experiencing is really a complexity crisis. Software is merely the most expedient way to implement complex systems. If you want to understand why writing reliable software is so hard, you are better off thinking of ecology than mathematics. It is the relationships between parts of a system that are most important, not the mathematical algorithms that make up the parts. We must realize that complicated systems are inherently less reliable than simpler ones, even if the software is totally bug free. The fundamental question about the high-tech baggage-handling system at the Denver Airport is not why it doesn't work but why it was attempted in the first place. Since the risk of these projects failing is so high, there must be a very significant benefit to be derived from them. The cost of the Denver system was $193 million. Would it be worth that much even if it worked?

Jeffrey M. Ratcliff
Orange, Calif.

Switching Brains

As an electrical engineer with over 20 years of experience in the design and implementation of analog electronics and servo-control systems, imagine my surprise when I found in "The Amateur Scientist," by John Iovine [Scientific American, October], that an ordinary differential amplifier in a closed-loop feedback-control system had been relabeled a neural network! The circuit described in the article is not a new concept based on the tenets of neural network theory but is in fact a widely used design with applications going as far back as World War II. The circuit and overall control loop of sensor, amplifier and motor drive can be rigorously and completely described by the standard methods utilized in closed-loop feedback-control system design.

Ronald B. Howes, Jr.
Minneapolis, Minn.

Iovine replies:

The fact that a 741 op-amp is being used in a closed-loop feedback-control system that has been described in electronics literature for the past 20 years in no way impedes the employment of such a system as a perfect example of an ideal neuron used in software or a hard-wired neuron used in neural network circuits. Hard-wired neurons configured in neural network systems were created using similar electrical designs as early as 1957, when Frank Rosenblatt built the successful Mark I Perceptron network. If we wired 100 or so op-amps in a neural configuration and trained it to play tic-tac-toe, would the resulting circuit be an electrical feedback system or a neural network?

Letters selected for publication may be edited for length and clarity. Unsolicited manuscripts and correspondence will not be returned or acknowledged unless accompanied by a stamped, self-addressed envelope.

50 AND 100 YEARS AGO

JANUARY 1945

"With the new Sperry Attitude Gyro, the airplane pilot can now be provided with an indication of the position of his plane even when the earth's surface cannot be seen. The instrument makes it possible to perform all aerobatic or acrobatic maneuvers without visual reference to the earth's surface."

"'The job of normal peace-time research is a private job, not a government job,' says Hon. Robert P. Patterson, Under Secretary of War. 'What the government may do, if it is called upon, is to furnish information and financial support. It may furnish counsel, even leadership. It must not, in the normal researches of peace, assume control.'"

"Progress in science as well as other branches of human endeavor depends upon the ability to communicate ideas to others by means of language. It is not required of the scientist that he be either an orator or a clever wielder of $64 words, only that he express himself clearly in everyday language."

"After nearly two years of laboratory tests and development, 'the strongest aluminum alloy yet available for commercial use' was recently announced by Reynolds Metals Company. The alloy, known as R303, is made with magnesium, zinc, and copper and has almost three times the compressive strength of structural steel."

"Perhaps the most interesting and promising of the proposed uses of glass-reinforced plastics are to be found in models for space-saving, structure-supporting, prefabricated kitchen and bathroom units. The two-sided assemblies, complete with full storage facilities, are intended to occupy a space only seven feet square, yet they are capable of supporting the entire structure of a house."

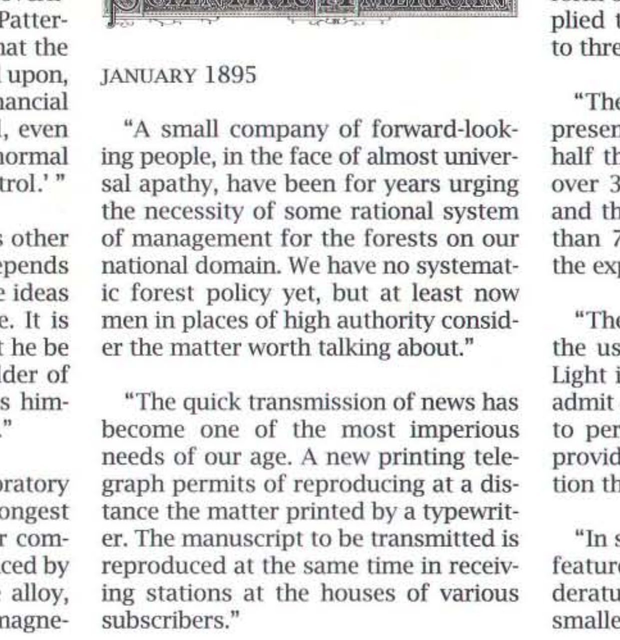
"The newest application of ion-exchange resins promises to be the purification of pectin and gelatin for use in the preparation of substitutes for blood plasma. Substitute blood plasma is being developed because of shortages, and present indications are that pectin and gelatin will serve as temporary and partial replacements."

"Too many old men are at the helm in science, which needs the originality of youth to keep pace with its opportunities for service to mankind."

JANUARY 1895

"A small company of forward-looking people, in the face of almost universal apathy, have been for years urging the necessity of some rational system of management for the forests on our national domain. We have no systematic forest policy yet, but at least now men in places of high authority consider the matter worth talking about."

"The quick transmission of news has become one of the most imperious needs of our age. A new printing telegraph permits of reproducing at a distance the matter printed by a typewriter. The manuscript to be transmitted is reproduced at the same time in receiving stations at the houses of various subscribers."



Ruth's combined sleeping and parlor car

"The 'American voice' has an unenviable reputation. It is apt to be shrill, strident, high-pitched, unmodulated. This quality adds an unnecessary aggravation to social life. It disorganizes the nerves, and increases the tendency to nervous prostration."

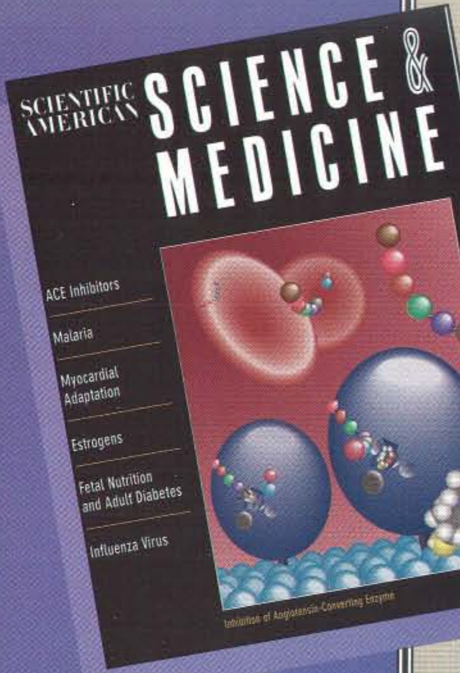
"Verily, the field of usefulness formerly held by the horse is narrowing daily. To steam, electricity, and the ubiquitous bicycle comes an ally in the form of explosive gas, so cunningly applied to the propulsion of vehicles as to threaten the horse's utter rout."

"There are in the United States at present 6,000,000 farms. About one-half the population of the republic or over 30,000,000 people live on them, and these farm dwellers furnish more than 74 percent of the total value of the exports of the country."

"The latest hygienic craze in Paris is the use of porous glass for windows. Light is freely admitted and the pores admit air. The minute holes are too fine to permit of any draught, while they provide a healthy, continuous ventilation through the apartment."

"In some things bigness is a valuable feature, in others, smallness is a desideratum. In the case of a battery, the smallest, lightest and most compact practicable battery made at present yields a large current (2 amperes) at a reasonably high voltage (1.1 volts). The battery consists of a zinc cell, closed with a hard rubber stopper, and containing an electrode formed of fused silver chloride."

"The combined sleeping and parlor car, shown in the accompanying illustration, depicts a notable feature recently patented by Mr. Linford F. Ruth. The cushions for the seats as well as for the bed are connected with the compressed air pipes of the train. The cushions are simply airtight bags of soft rubber or other suitable material and can be inflated by opening valves in connecting pipes, or be collapsed and compactly stored."



SCIENTIFIC AMERICAN
SCIENCE & MEDICINE
covers it all…

Gastrointestinal Motility
Lawrence R. Schiller
Baylor University Medical Center

Confocal Laser Microscopy
Edwin de Feijter
Meridian Instruments, Inc.
Okemos, Michigan

Taxol
Susan Band Horwitz
Albert Einstein College of Medicine, New York

Pathogenesis of Sepsis
Michel P. Glauser
Centre Hospitalier Universitaire Vaudois, Lausanne, France

Cerebral Ischemia
William Pulsinelli
University of Tennessee Health Science Center

Second-Generation HIV Vaccine
Emilio A. Emini
Merck Research Laboratories
West Point, Pennsylvania

…and more, including:

RNA Editing

Cryosurgery

Barrett's Esophagus

LET THE EXPERTS HELP YOU UNDERSTAND CLINICAL RESEARCH AND ITS PRACTICAL APPLICATIONS IN…

NEW

SCIENTIFIC AMERICAN
SCIENCE & MEDICINE

Introducing SCIENTIFIC AMERICAN **SCIENCE & MEDICINE**, an exciting, new magazine for the medical community. From medical student to experienced practitioner, everyone in clinical medicine needs to keep abreast of the vast quantity of scientific research being conducted all over the world and how the findings can impact their work.

You'll find four-color graphics and a wide variety of topics covered in each bi-monthly issue of SCIENTIFIC AMERICAN **SCIENCE & MEDICINE**. Published in the high editorial standards of SCIENTIFIC AMERICAN, this new medical magazine will soon become an indispensable part of your regular professional reading. With articles written by leading authorities actively engaged in scientific work, you'll not only get up-to-the-minute information on headline topics in the medical field, but you'll also get a sense of where medical science is headed. You can also stay further informed with SCIENTIFIC AMERICAN **SCIENCE & MEDICINE** departments: About Benefits and Costs, AIDS Day by Day, SCIENTIFIC AMERICAN MEDICINE Updates, Patents Column and bioscience glossary.

Subscribe now! Simply fill out your Subscription Form below, put it in an envelope and mail today!

Welcome to SCIENTIFIC AMERICAN **SCIENCE & MEDICINE.**

SUBSCRIPTION FORM

YES! Please enter my subscription to SCIENTIFIC AMERICAN **SCIENCE & MEDICINE**.
I understand my subscription is guaranteed by SCIENTIFIC AMERICAN:
I may cancel at any time for a refund on the undelivered portion.

☐ Two-year subscription (12 issues) US$121 ☐ One-year subscription (6 issues) US$75
☐ AIR RATE Two-year subscription US$209 ☐ AIR RATE One-year subscription US$109
☐ Check enclosed ☐ Bill me
☐ Charge my: ☐ VISA ☐ MasterCard
☐ Eurocard ☐ American Express

CARD NUMBER

EXPIRES

SIGNATURE

NAME

ADDRESS

CITY/PROVINCE POSTCODE COUNTRY

P5AX76

Remit US Funds drawn on a US bank.
Addresses within the U.S.: US$89 (2yrs), US$59 (1yr).
Enclose this form in an envelope & mail to:
SCIENTIFIC AMERICAN SCIENCE & MEDICINE, P.O Box 3182, Harlan, IA 51593-0373

SCIENCE AND THE CITIZEN

For Whom the Bell Curve Really Tolls

A tendentious tome abuses science to promote far-right policies

Rarely do 800-page books that are crammed with graphs reach best-seller lists. *The Bell Curve,* an inflammatory treatise about class, intelligence and race by the late Richard J. Herrnstein, a psychology professor at Harvard University who died last September, and political scientist Charles Murray of the American Enterprise Institute, is an exception. The book's deeply pessimistic analysis of U.S. social woes, together with its conservative policy prescriptions, has hit a nerve.

JEFFRY MYERS FPG International

EDUCATION can benefit all, a truth being forgotten in the clamor over The Bell Curve.

Publishing *The Bell Curve* may have been a calculated political move on the part of its authors. As the country lurches to the right, many people will be seduced by the text's academic trappings and scientific tone into believing its arguments and political inferences well supported. Those readers should think again.

The Bell Curve depicts a frightening future in which, absent strong corrective measures, a "cognitive elite" will live in guarded enclaves distant from the dull masses. Opportunities for the underclass will become limited as tolerance evaporates. Strict policing will be widely accepted, and racial hostility will most likely spread. The least intelligent denizens of this dystopia will be consigned to a "high-tech and more lavish version of the Indian reservation."

This apocalyptic vision is presented as the consequence of unpalatable, undeniable "facts" about inheritance and intelligence. But the thesis rests on curiously twisted logic. Its authors have been highly selective in the evidence they present and in their interpretation of ambiguous statistics. The work is "a string of half-truths," states Christopher Jencks, a sociologist at Northwestern University.

The arguments stem from the same tradition of biological determinism that led, not so long ago, to compulsory sterilizations in the U.S. and genocide elsewhere. The notion is that individuals' characteristics are both essentially fixed by inheritance and immune to alteration by the environment. Efforts to help those who are unfortunate by reason of their genes are unlikely to be rewarded. Solutions, therefore, should include those Murray has long advocated: abolish welfare, reduce affirmative action and simplify criminal law.

Herrnstein and Murray produce data suggesting that intelligence—as assessed by a high IQ score—is increasingly important to economic success. They also argue that people who have low scores—including disproportionate numbers of blacks—are more likely than others are to fall prey to social ills. The two accept evidence from studies of twins reared apart that there is a large heritable component to IQ scores: they estimate it to be 60 percent. The writers declare themselves agnostic on the question of whether racial differences in IQ scores are genetic, although they are clearly inclined to favor that possibility.

Herrnstein and Murray concede that just because a trait has a heritable origin does not mean it is unchangeable. Nearsightedness is one example of an inherited, modifiable condition. But they decide, on the basis of a questionable look at the data, that "an inexpensive, reliable method of raising IQ is not available." This conclusion is used to justify an attack on programs aimed at helping society's most vulnerable; the authors prefer to let the genetically disadvantaged find their own level. Evidence that does not accord with Herrnstein and Murray's way of thinking—such as the observation that IQ scores worldwide are slowly increasing—is acknowledged, then ignored.

Leaving aside the substantial and unresolved issue of whether a single number can adequately summarize mental performance, *The Bell Curve* plays fast and loose with statistics in several ways. According to Arthur Goldberger, an econometrician at the University of Wisconsin who has studied genetics and IQ, the book exaggerates the ability of IQ to predict job performance. Herrnstein and Murray assert that scores have an impressive "validity" of about 0.4 in such predictions. They report that the Armed Forces Qualification Test, an IQ surrogate, has a validity of 0.62 at anticipating the success of training for mechanical jobs. Yet many of the measures used to assess validity include supervisors' ratings, which are subject to bias, Goldberger notes. Furthermore, the validities that the duo see as so revealing are, in fact, hypothetical quantities that no employer would expect to find in prospective employees. "It's really bad stuff," Goldberger says.

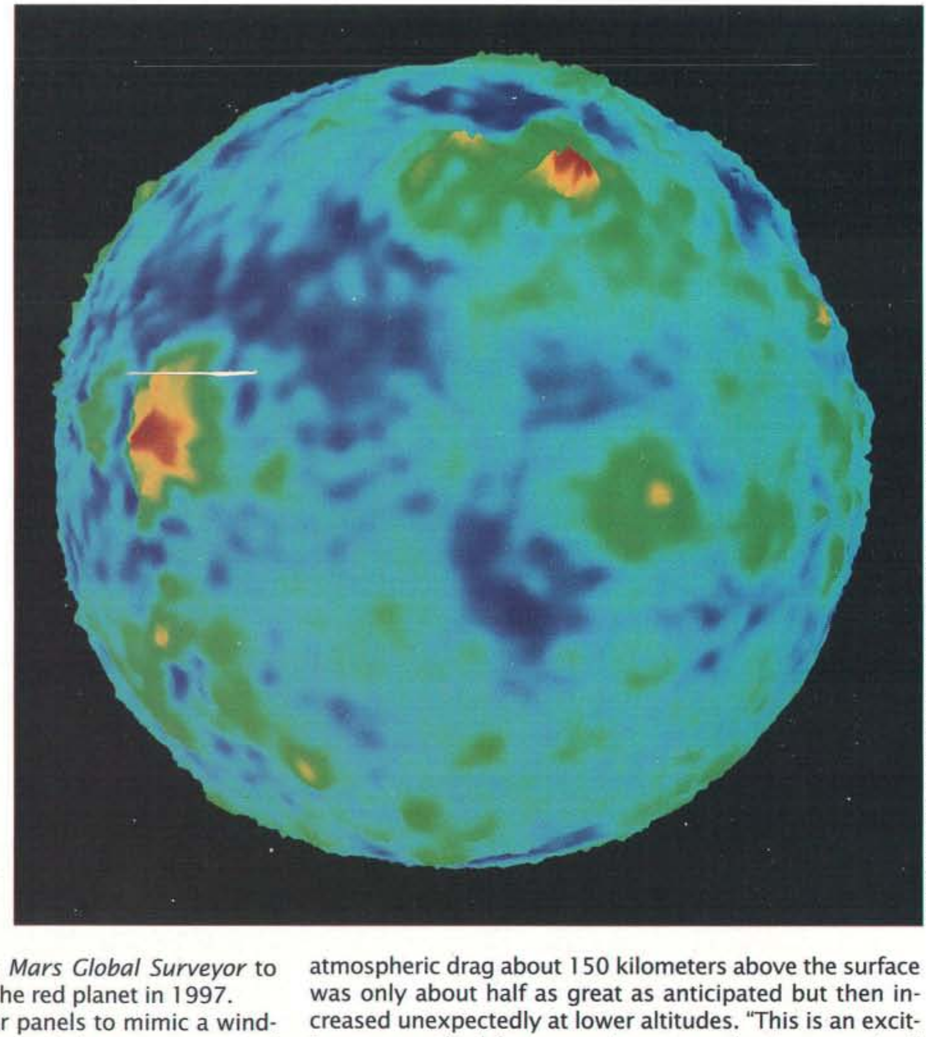
Other correlations that the writers establish between social ills and low IQ scores are equally suspect. Herrnstein and Murray put great weight on comparisons between the ability of IQ scores and parental socioeconomic status to predict what will happen to young people. Yet the measures of socioeconomic status they use cannot ensure that homes are equally stimulating. The point is crucial because numerous studies have demonstrated that early childhood surroundings have a large role in molding IQ scores—while very few studies have indicated a significant role for heredity. Consequently, conclusions about the dominance of IQ cannot be taken at face value. Leon Kamin, a psychologist at Northeastern University and well-known critic of research on intelligence, maintains that interactions between genes and environment make attempts to weigh nature against nurture "meaningless."

Herrnstein and Murray's hereditarian bias is also obvious in their account of a study of 100 children from varying ethnic backgrounds who were adopted into white families. The study got under way in the 1970s. At age seven, the black and interracial children scored an average of 106 on IQ tests—considerably better than the national average of black children and close to levels scored by white children. A decade later researchers Sandra Scarr of the University of Virginia and Richard A. Weinberg of the University of Minnesota found that the IQs of the black and interracial children had declined to 89 and 99, respectively, whereas those of white adoptees had fallen from 112 to 106.

Scarr and Weinberg concluded that racially based discrimination at school probably explained the drop in the black youngsters' scores. Jencks agrees: "The results are perfectly consistent with the difference being due to something in the early home environment and, for older kids, their experience in school." But Herrnstein and Murray interpret the findings differently: "Whatever the environmental impact may

Deathbed Revelations

The *Magellan* spacecraft, which produced spectacular radar images of the surface of Venus, gave its life to science when it plunged into that planet's murky atmosphere on October 12. Project scientists had maneuvered *Magellan* into a low, and ultimately sacrificial, orbit so that it could map Venus's gravitational field. Tiny wiggles in the orbit betrayed local variations in the mass of the planet, clues to its internal structure. The resulting gravity map is shown here superimposed on an exaggerated-relief image of Venus's topography. Gravitational highs are rendered in red; gravitational lows are displayed in blue.



JET PROPULSION LABORATORY/NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

As *Magellan* dipped closer to its infernal doom, it performed unprecedented acrobatic feats. The drag created as the craft sped through the thin upper atmosphere pulled it ever downward, producing the first real-world test of aerobraking. The new fuel-saving technique will be used by the *Mars Global Surveyor* to help guide it into orbit around the red planet in 1997.

Magellan also turned its solar panels to mimic a windmill. Technicians measured how much thrust was required to keep the probe from spinning—information that yielded surprising data about Venus's atmosphere. According to Robert H. Tolson of George Washington University, the atmospheric drag about 150 kilometers above the surface was only about half as great as anticipated but then increased unexpectedly at lower altitudes. "This is an exciting new method for measuring atmospheric properties," he says, one that may soon be applied to earth-orbiting satellites. *Magellan* may live no more, but new insights and questions have arisen from its ashes. —*Corey S. Powell* have been, it cannot have been large."

The Bell Curve's most egregious failing, however, may be its bleak assessment of educational efforts to improve the intellectual performance of children from deprived backgrounds. Herrnstein and Murray cast a jaundiced eye over Head Start and other efforts for at-risk youngsters—projects that have been claimed to produce long-lasting gains in IQ, a possibility that would not square well with biological determinism. Herrnstein and Murray downplay such results, noting that such interventions are too expensive to be widely used. The only one they are enthusiastic about is adoption, which, paradoxically, they accept as having a positive effect on IQ. "Their treatment of intervention wouldn't be accepted by an academic journal—it's that bad," exclaims Richard Nisbett, a psychology professor at the University of Michigan. "I'm distressed by the extent to which people assume [Murray] is playing by the rules."

Jencks is also unhappy with the book's conclusions about education. "Herrnstein and Murray are saying Head Start didn't have a profound effect. But that doesn't tell us that we couldn't do a lot better if we had a different society," he says. "In Japan, for example, children learn more math than they do in the U.S. because everybody there agrees math is important."

Scarr, who accepts a substantial role for heredity in individual IQ differences, insists that efforts to boost intellectual functioning in disadvantaged youth can deliver results. "There's no question that rescuing children from desperately awful circumstances will improve their performance," she notes.

Scarr also points out that ameliorating a child's environment may reduce social problems, regardless of its effect on IQ. "The low-IQ group deserves a lot more support than it is getting," she argues. "Other societies manage not to have the same levels of social ills as we do." Edward F. Zigler, a prominent educational psychologist at Yale University, asserts that "in terms of everyday social competence, we have overwhelming evidence that high-quality early education is beneficial."

Therein lies the fatal flaw in Herrnstein and Murray's harsh reasoning. Even though boosting IQ scores may be difficult and expensive, providing education can help individuals in other ways. That fact, not IQ scores, is what policy should be concerned with. *The Bell Curve*'s fixation on IQ as the best statistical predictor of a life's fortunes is a myopic one. Science does not deny the benefits of a nurturing environment and a helping hand. *—Tim Beardsley*

The Great Attractors

Chemical guides direct young neurons to their final destinations

More than a century ago the renowned Spanish neurobiologist Santiago Ramón y Cajal discovered the growth cone, "that fantastic ending of the growing axon." His find partially explained one of the most fundamental and dynamic events in embryonic development. These "living battering rams," as he observed, sprout from nerve cells and forge ahead toward select tissues. Hence, he suggested that these structures enable young neurons to wire the synaptic links that form an adult nervous system. Until recently, though, no one had figured out how the growth cones know where to go.

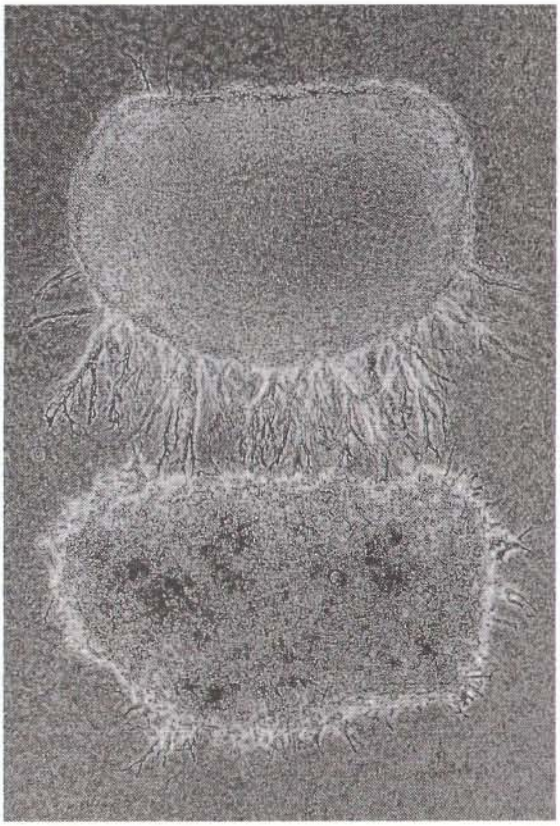
Cajal himself, it turns out, had the right idea. He proposed that the target tissues might release certain diffusible chemicals that, like a trail of bread crumbs, could lure the advancing axons from afar. Following this path, a team led by Marc Tessier-Lavigne of the University of California at San Francisco and the Howard Hughes Medical Institute identified two such chemotropic proteins. It has christened them netrin-1 and netrin-2, after the Sanskrit *netr*, meaning "one who guides."

Both proteins promote and orient the growth of so-called commissural axons in the developing spinal cord of chickens and rodents. These axons branch from nerve cells in the dorsal spinal cord and travel around its circumference to tissues in the front known as the floor plate. From there, they turn toward the brain. Studies done in vitro have shown that a collection of floor-plate cells can elicit axonal outgrowth of this kind from dorsal spinal-cord explants. Nevertheless, because the floor plate is so small, workers had been unable to isolate its active ingredients.

Tessier-Lavigne and his colleagues managed to avoid that problem altogether. They compared the floor plate's allure with that of more accessible tissues and found that the cell membranes in a developing chick brain could also draw commissural axons at a distance. The team purified the netrins from some 25,000 chick brains. To confirm that these proteins were indeed the spinal cord's chemical bait, the group introduced netrin-1 RNA into a line of mammalian cells. These custom-made cells then produced netrin-1 and attracted axons as floor-plate cells would.

Although both netrin-1 and netrin-2 were present in the chick membrane, floor-plate cells make only netrin-1. "The netrin-1 transcript is expressed at high levels in the floor plate," Tessier-Lavigne says, "whereas netrin-2 is expressed at lower levels over the ventral two thirds of the spinal cord." He speculates that this distribution might explain the path commissural axons typically take. Because higher levels of netrin-1 linger near the floor plate, the outgrowing axons most likely travel toward an ever increasing amount of netrin to reach their destination.

As further evidence that the netrins govern this growth, the same pattern of circumferential migration seems to occur in other species. The researchers have discovered that the netrins resemble unc-6, a protein that guides the growth of certain axons in a nematode. And Corey S. Goodman's laboratory at



MARC TESSIER-LAVIGNE *University of California at San Francisco*

BATTERING RAMS—or growth cones from commissural axons—are lured toward floor-plate cells by chemical cues.

the University of California at Berkeley recently isolated a netrin gene in the fruit fly *Drosophila.* "This shows a strong conservation of biological functions between species," Tessier-Lavigne notes.

The researchers studying worms, flies and vertebrates plan to collaborate extensively. They are now testing whether netrins and unc-6 function equally well in vertebrate and invertebrate systems. "Aside from axonal projections, unc-6 controls the circumferential migrations of cells in worms," Tessier-Lavigne adds. "So cell migrations in vertebrate embryos might use netrins as guidance cues."

Such analogies may help the group answer other questions as well. The scientists have uncovered a slightly smaller protein, dubbed NSA for netrin-synergizing activity, that seems to influence netrin potency. Perhaps NSA, like certain proteins in other signaling pathways, mediates how well the netrins bind to their receptors. "We really want to know if NSA is an essential cofactor or not," Tessier-Lavigne states. Also, the netrins tend to adhere to cell surfaces for reasons as yet unknown.

What is known is that these novel proteins are probably just two words in an entire language of chemical instructions that direct embryonic development. Neurobiologists hope to discover chemicals that can ward off outgrowing axons and thus prevent them from making faulty connections. Perhaps other kinds of cues exist as well. At any rate, Tessier-Lavigne predicts that progress will be swift because similar chemical words seem to speak of the same biological functions in different species. "Now we can go back and forth between different systems and share our insights," he says. For a while, it seems the netrins will bring scientists together as surely as they connect searching axons. *—Kristin Leutwyler*

Socializing with Non-Naked Mole Rats

RAYMOND MENDEZ *Work as Play*

Big and hairy, the Damaraland mole rat is not as renowned as its hairless cousin. Nevertheless, this species has proved just as intriguing as the naked mole rat of zoo and cartoon fame. Both forms of mole rat are eusocial—that is, they live in groups in which only a queen and several males reproduce, whereas the rest of the colony cooperates to care for the young. This behavior—like that of termites and ants—is found in very few mammals, and it has remained a puzzle of natural selection.

By comparing Damaraland and naked mole rats, Jennifer U. M. Jarvis and Nigel C. Bennett of the University of Cape Town and others have begun to determine the characteristics that appear central to the evolution of eusociality—and hair is clearly not one of them. "The Damaraland is important because it does not have many of the characteristics of the naked mole rat," notes Paul W. Sherman of Cornell University. So it "tells us something that we did not know." The degree of genetic relatedness between members of a colony, for instance, does not appear as crucial to eusociality as some had believed. In the case of naked mole rats, siblings raise one another because the survival of a sister or brother is virtual cloning. A Damaraland colony, however, appears much more genetically diverse. Once a queen dies, these mole rats wait to reproduce until another female is introduced from somewhere else—at least in the laboratory.

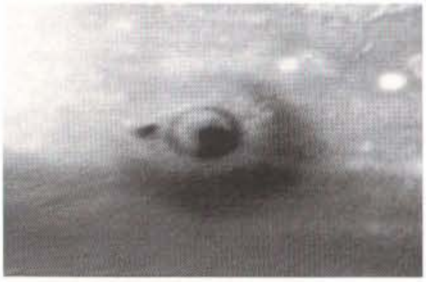
Instead ecological determinants seem more significant to eusociality. Both Damaraland and naked mole rats live in arid regions where the food supply, underground tubers, is sparse and rainfall unpredictable. Cooperative living ensures finding these precious resources—solitary animals would be unable to tunnel extensively enough to locate adequate sustenance. The other, noneusocial forms of mole rats live in regions where food is more readily available. In other words, the more patchy the food, the more the cooperation. *—Marguerite Holloway*

Picking Up the Pieces

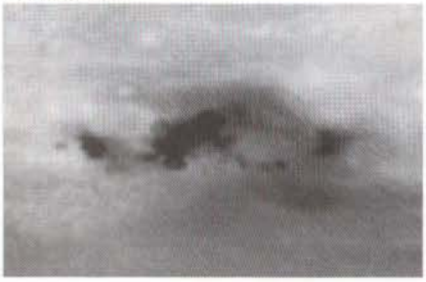
Astronomers mull over the lessons of the great comet crash

Last Halloween hundreds of astronomers gathered in Bethesda, Md., to play an unusual kind of trick or treat. The treat was a bumper crop of observations about the death of Shoemaker-Levy 9, the comet that smashed into Jupiter last July. The trick was trying to make sense of it all. And the costume of choice was an "I survived the comet crash" T-shirt.

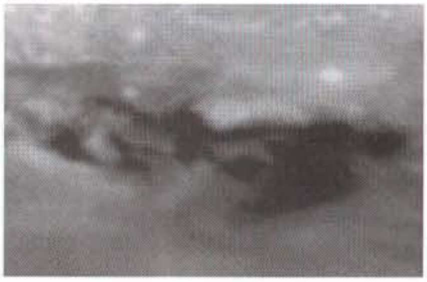
JULY 18, 1994



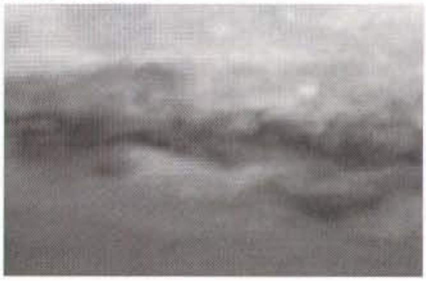
JULY 23, 1994



JULY 30, 1994



AUGUST 24, 1994



HEIDI HAMMEL *Massachusetts Institute of Technology and NASA*

***DARK RINGS** of dusty material formed around the impact sites of the larger fragments from Shoemaker-Levy 9. As the weeks passed, the material was dispersed by east-west flows in Jupiter's stratosphere.*

Although many of the scientists at the Division of Planetary Sciences meeting were seeing their colleagues' results for the first time, some indications of a consensus began to emerge. For instance, researchers are starting to home in on the exact nature of Shoemaker-Levy 9. Initial analyses of the debris kicked up by the impacts had shown no sign of water. Because normal comets are thought to contain ice, some researchers had proposed that Shoemaker-Levy 9 might actually have been a fragmented asteroid.

But G. L. Bjoraker of the National Aeronautics and Space Administration Goddard Space Flight Center reported that he saw distinct evidence of water, in a quantity equivalent to a kilometer-wide ball of ice. "There's no deficiency of ice in that comet," declared Eugene M. Shoemaker of Lowell Observatory, satisfied that the object bearing his name would not be downgraded to asteroid status. Other researchers detected a complex melange of Jupiter and bits of vaporized comet—metals, carbon monoxide and a great deal of sulfur.

Astronomers are also beginning to arrive at a plausible explanation for the unexpected, vast, dark splotches that appeared on Jupiter after the crash. Carl Sagan of Cornell University argued that the patches probably consisted of carbon compounds derived from organic molecules in the comet. By the end of the day, most of his colleagues seemed to agree.

Other aspects of the crash, however, eluded explanation. The bright flashes and gargantuan plumes were the source of much debate. Heidi B. Hammel of the Massachusetts Institute of Technology presented evidence from the *Hubble Space Telescope* that all the plumes rose about 3,000 kilometers above the top of Jupiter's clouds—even though the fragments ranged greatly in size. Why should a small collision make as tall a splash as a big one? "You'll have to ask the modelers," she shrugged. "But it's true."

Furthermore, every researcher seemed to detect a slightly different pattern of flashes. Imke de Pater of the University of California at Berkeley described observations made using the giant Keck telescope in Hawaii that showed that one piece of comet produced three flare-ups. She theorized that the fragment disintegrated into a stream of rubble 1,800 kilometers long. The initial flare represented the first bits of comet hitting the far side of the planet. The second flare appeared when the resulting fireball rose into direct view. The final, brightest flash occurred when material fell back into the atmosphere, heating it to an incandescent glow.

De Pater's analysis resulted from a consultation with Mordecai-Mark Mac Low of the University of Chicago, who presented computer simulations of the impact. Mac Low fielded many questions, but the astronomers in attendance seemed generally to accept his model. Shoemaker remained skeptical. "It's nonsense," he exclaimed, arguing that the returning material would not impart enough energy to cause intense flashes. A better understanding of the radiation pulses will clarify how energy spreads after huge impacts—including the ones that may have caused mass extinctions on the earth.

The many remaining mysteries were accompanied by some magic. Hammel wowed the meeting with images of the dark splotches, many of which initially formed delicate rings and improbable "mustache" shapes. Andrew P. Ingersoll of the California Institute of Technology proposed that the complicated shapes resulted from material thrown up by the impact and from an associated gravity wave, a kind of up-and-down atmospheric disturbance. The waves presumably caused vaporized organic material to condense high in Jupiter's atmosphere. Hammel estimated that the initial "hole" in the atmosphere that had created the rings was some 500 kilometers wide—about the size of Texas.

The day's $64,000 question was, of course, how often giant comets are wont to crash into Jupiter—and, by extension, the earth. Chains of craters on Jupiter's moons appear to record the impacts of earlier comets, which, like Shoemaker-Levy 9, broke up into smaller pieces. Paul M. Schenk of the Lunar and Planetary Institute in Houston analyzed those craters and concluded that comets plow into the Jovian system remarkably often: every 150 years or so.

The lack of final answers did not faze anyone—Shoemaker least of all. "It was great; I was in heaven," he summed up after the scientific sessions. "Now I want the grand synthesis." Ingersoll also encouraged an intensive effort to understand the observations because "this may happen again." He was thinking mainly about another event on Jupiter: his calculations indicate that the planet is at least 10,000 times (perhaps even a million times) as likely to be hit as the earth is. But his comment recalled the disconcerting possibility that nature's next Halloween treat could land a little closer to home. *—Corey S. Powell*

Mystery of the Missing Dynamo

Astronomers cannot explain the galaxy's magnetic field

A planet, a star and a galaxy may not look much alike. Yet each rotates and has a magnetic field. For decades, astrophysicists have held that the two attributes are related: magnetic fields are generated by rotating charged particles, through a process called a dynamo. Now an increasingly vocal assortment of theorists is arguing that for the sun and the Milky Way—and by extension, for all stars and galaxies—the usual dynamo mechanism falls apart. "It would take 10^{26} years to create the Milky Way's magnetic field," notes Russell M. Kulsrud of Princeton University. The galaxy has been around for at most 10^{10} years.

"Fundamental dynamo theory is in big trouble," agrees Patrick H. Diamond of the University of California at San Diego. And if the dynamo cannot be repaired, astrophysicists will have no way of understanding how stars and galaxies come to have their magnetic fields.

Trouble is not new to dynamos. At its root is the tortuous behavior of a magnetic-field line when it is embedded in a highly conducting fluid, such as that in a star or the interstellar medium. Charged particles spin around the line, trapping and forcing it to share the motions of the plasma. But in 1934 Thomas G. Cowling of the University of Leeds in England proved that simple, symmetrical fluid flows cannot generate magnetic fields. The curse of this antidynamo theorem was lifted only in 1955, when Eugene N. Parker of the University of Chicago described how turbulence in the hot, ionized atmosphere of the sun might amplify a small primordial field.

Parker's model, honed by Max Steenbeck and others at Potsdam University in Germany, became the favorite of astrophysicists. Now researchers are sighing that it, too, is doomed—by the very turbulence that saves it from the antidynamo theorem.

Consider, for example, how the dynamo works in a galaxy. The galactic disk rotates more slowly at its outer edges. A field line that lies along a radius will be dragged around with the disk. Yet the line will trail at the edge and eventually get wound into a tight spiral. The one line then passes repeatedly through each small region of plasma, mimicking many field lines and yielding a strong overall field.

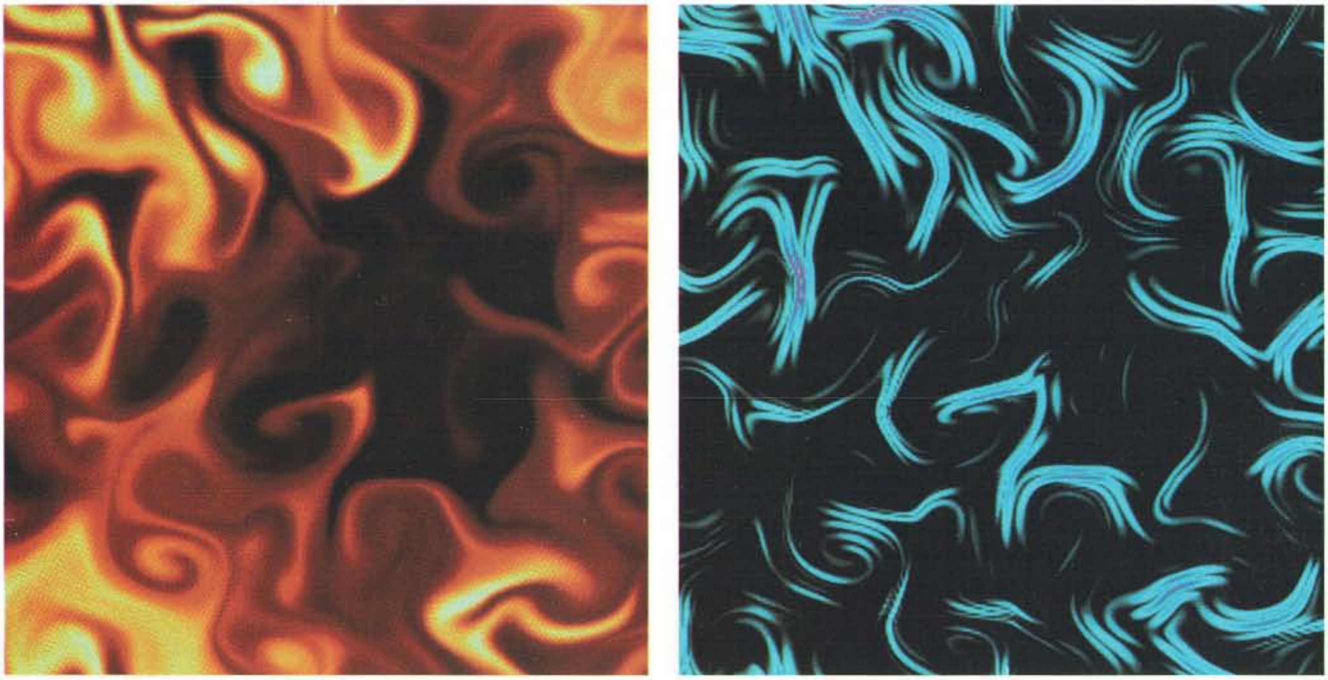
If a small loop of line happens to jut out of the plane of the galaxy, Coriolis forces (which determine the direction in which tornadoes twist) curl it in opposite directions on either side of the disk. Many such curls line up into ringlets, creating magnetic fields along the axis of rotation. The twists, however, resist bending and tend to uncurl. To release tension without losing the axial field, the loops must migrate toward one another, break up and rejoin into larger loops.

Here is the crunch. The field lines are trapped in the fluid, to an extent measured by the magnetic Reynolds number. This number increases with the electrical conductivity of the plasma and the physical extent of the magnetic field. In the earth's molten core the Reynolds number is about 100. In a star or galaxy where distances are enormous, Reynolds numbers of 10^{10} or 10^{19}, respectively, ensure that trapping is almost complete.

The curls may, however, diffuse and recombine if they are very small, so that effectively their Reynolds number is low. In a series of papers, Kulsrud has shown that the twists would then have to be so tightly wound that they would pull back, stop winding and halt the dynamo before it had generated any overall field for a galaxy.

Samuel I. Vainshtein of the University of Chicago argues further that the twists are fractal [*see illustration below*]. Extending to extremely small scales, they are even stiffer. "These parasites eat up all the energy and contribute nothing to the field," he chuckles. (They also ate up another popular dynamo invented by Vainshtein and Yakov B. Zel'dovich in 1972. In this model a magnetic-field loop is stretched out, twisted and folded over repeatedly, thereby amplifying a small initial field.)

Vainshtein recalls presenting his preliminary ideas two years ago at a meeting held in Cambridge, England. "The astrophysical community was mad. They said, 'What about the sun? It works there!' " But the solar dynamo, it



***FRACTAL LOOPS** of magnetic fields* **(left)** *and electric current* **(right)** *absorb all the energy in an astrophysical dynamo, bringing it to a grinding halt. This simulation was provided by Fausto Cattaneo of the University of Chicago.*

turns out, has some troubles of its own.

The galactic dynamo, after all, has one small success. Philipp P. Kronberg of the University of Toronto, Richard Wielebinski of the Max Planck Institute for Radioastronomy and others have found that the magnetic-field lines of the Milky Way lie along its spiral arms—wrapped around just as the dynamos would have it. Solar models, however, while explaining beautifully the 22-year sunspot cycle, were not so lucky. In the late 1980s helioseismological data revealed that near its poles, the sun's rotation is slower at deeper layers. The models required the inner rotation to be faster.

Theorists have been quick to come to the dynamo's rescue. Parker speculates that the fast-moving gas of cosmic rays permeating the galaxy could blow out giant magnetic bubbles, which easily reconnect outside the disk. George B. Field of Harvard University protests Kulsrud's calculations, saying supernova explosions would greatly alter the energetics of the turbulent field. Or, as Ellen G. Zweibel of the University of Colorado points out, collisions between charged ions and neutral atoms in the plasma could help the field lines diffuse faster. The nuts and bolts of the sun's dynamo could be hidden underneath the convective zone, as yet beyond the reach of sunquake studies.

Few of these proposals are backed up with detailed calculations. The only other explanation for astrophysical magnetic fields is that they were created in the early universe. Still, these fields would extend no farther than the universe did at the time—far too short to be galactic. Besides, no mechanisms for generating such strong fields are known. Small "seed" fields, which a dynamo might amplify, are easier to come by.

Indeed, the earth appears to possess the only dynamo that is not fatally flawed. Tornadoes in the molten core are said to create its dipole field, although questions persist about why the field sporadically flips. "I'm not an optimistic fellow," declares Paul H. Roberts of the University of California at Los Angeles. "But I think we'll get it right by the end of the century." Experimentally, fluid dynamos have not been demonstrated to work—apart from a dynamo (undesired) that reportedly started up in a nuclear reactor in Beloyarsk in the former Soviet Union.

The geophysicist Walter Elsasser once related to a friend Cowling's attempt to make a dynamo. "If that simple idea does not work," remarked Albert Einstein, "then dynamo theory will not work either." *—Madhusree Mukerjee*

Holes in Ozone Science

Researchers look at loss of the protective layer above our heads

Concerns about the thinning ozone layer over Antarctica are familiar to most people. But part of the story is not so well known: the thinning in midlatitude regions, which include some of the most populated areas on the globe. And although the mechanism behind ozone depletion at the poles is generally understood, details of the process at midlatitudes remain obscure.

Recently, however, a team of researchers has begun to characterize the cycle of ozone loss and regeneration taking place right over our heads in the Northern Hemisphere. The new findings describe how several compounds, not solely the infamous chlorofluorocarbons (CFCs), contribute to such depletion. The results have significant implications for a U.S. plan to build supersonic commercial aircraft.

Scientists and policymakers first considered midlatitude ozone depletion in the early 1970s. They examined the environmental impact of a fleet of airplanes that would fly faster than the speed of sound at altitudes ranging from 17 to 20 kilometers (the lower stratosphere). With few measurements to go on, most experts suspected that nitrogen oxides from the exhaust would prove to be the most significant destroyer of ozone.

TOM REDDY NASA Ames Research Center

PREPARING THE PLANE for its flight, scientists load equipment into the craft's nose. The instruments will measure levels of hydrogen oxides, which account for up to 50 percent of ozone loss in the lower stratosphere over midlatitude regions.

As it turned out, severe ozone depletion first surfaced in another corner of the planet. In 1985 the British Antarctic Survey discovered an ozone hole over the South Pole, and atmospheric scientists scrambled to determine what had produced it. By 1991 CFCs had been firmly established as the principal cause. These compounds contain chlorine, which can break apart the ozone molecule, O_3. This triad of oxygen absorbs ultraviolet (UV) light, using the energy to break O_3 into O and O_2 and preventing harmful radiation from reaching the earth. CFCs and other chemicals prevent O and O_2 from recombining and replenishing the ozone layer.

The world continues to keep a close eye on the Southern stratosphere. In 1993 ozone levels there dropped by around 60 percent. Perhaps because it is not characterized by a gaping hole, midlatitude thinning has received less attention. Nevertheless, observations may point to a trend: between 1978 and 1990, ozone levels over North America dropped by 0.5 percent per year. In 1993 the total loss reached 7.5 percent.

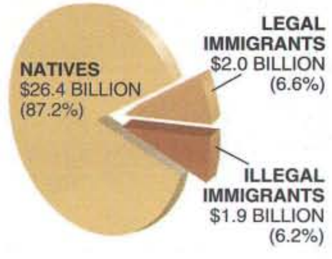
To determine what underlies depletion in these regions, the National Aeronautics and Space Administration's Stratospheric Photochemistry, Aerosols and Dynamics Expedition Project sent nine flights into the midlatitude stratosphere during 1993. Scientists believe warmer temperatures, variable patterns of air movement and the presence of sulfuric acid aerosol particles may differentiate midlatitude ozone loss from polar processes. So NASA's program measured concentrations of ozone, aerosol particles, nitrogen oxides, hydrogen oxides and CFCs, among other compounds.

According to Paul O. Wennberg of Harvard University, it appears that hydrogen oxides—not nitrogen oxides—account for up to 50 percent of the ozone loss observed below 20 kilometers. Recent models had predicted this finding, but no one could measure hydrogen oxides in the lower stratosphere, because they are present in such low concentrations. To address this problem, the group, under the direction of James G. Anderson, developed a solid-state laser device able to record concentrations below 0.1 part per trillion.

The data further "show that there is a region where it might be safe to fly supersonic aircraft," where nitrogen oxides from exhaust should not be detrimental, explains Harold S. Johnston of the University of California at Berkeley. What defines such a safe flying zone, however, seems to shift. New knowledge has brought the top of such a path up to an altitude of 20 kilometers from the 13 kilometers cited in the 1970s. Johnston cautions that nitrogen oxides are still believed to be the most important ozone-destroying compounds above 30 kilometers.

Fear and Self-Loathing in America

The U.S. is a nation of immigrants, but newcomers are not very popular these days—the passage of California's Proposition 187 is but one example (the law bars illegal immigrants from medical care and schooling). Immigrants are seen by many as usurping scarce jobs and draining government funds. In reality, according to recent studies by the Urban Institute in Washington, D.C., this perception is skewed. Data show that immigrants create more jobs than they fill. In 1989 total immigrant income was $285 billion, about 8 percent of all reported income (immigrants make up 7.9 percent of the population). Much of this money is spent on U.S. goods and services. As for public assistance, the share utilized by illegal immigrants is relatively small (*right*). *—Marguerite Holloway*



The economic and political pressures to understand this particular chemistry are vast. A phalanx of, say, 500 high-speed planes could bring the U.S. airline industry $100 billion in sales. For now, the supersonic jets must remain on the drawing board. Richard S. Stolarski of the NASA Goddard Space Flight Center points out that despite recent findings, extensive analysis of the airplanes' environmental and economic viability is still necessary.

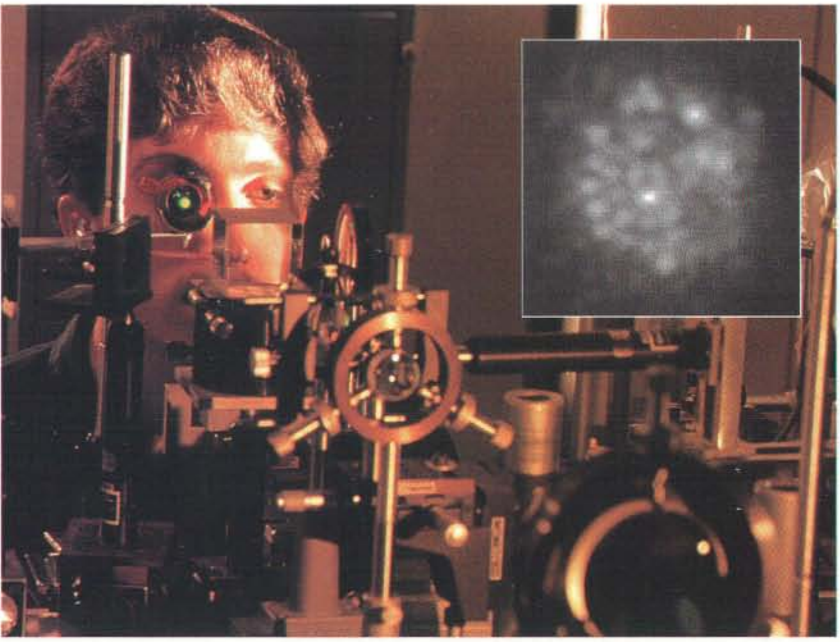
The health effects of ozone depletion must also be considered. Reports of increased UV radiation reaching North America have been linked to falling stratospheric ozone levels. Researchers continue to study the possible long-term effects on plants and animals: exposure to high doses contributes to skin cancer and can weaken the immune system. And the links are not always straightforward. A Canadian team led by Max L. Bothwell of the National Hydrology Research Institute in Saskatchewan reported last summer that higher UV exposure resulted in larger populations of algae. Apparently, the radiation harmed the insects grazing on the algae more than it damaged the algae. As Bothwell says, "The effects of UV radiation are more complex than we thought."

Indeed, the entire issue of ozone depletion over midlatitudes continues to reveal unexpected complexities. Anderson sums up the problem faced by everyone studying such depletion with one word: uncertainty. "We just do not understand the midlatitude stratosphere from top to bottom," he says. "Clearly, there is the potential for further surprise." *—Sasha Nemecek*

Seeing the Cells That See

Ever since the eye's rods and cones were discovered, scientists have been trying to observe them alive and in action. But the retinal photoreceptors, which change light into electrical signals the brain can process, are so tiny and their flashes of activity so brief that they have eluded researchers.

Finally, last fall, a team led by David R. Williams of the University of Rochester managed to peek at and photograph human cones. As demonstrated in this picture, researchers used a laser to illuminate the retina; a high-resolution camera, like those astronomers use, recorded the image. The cones, shown here in the black-and-white inset, are three microns wide and are responsible for color and daytime vision. *—Marguerite Holloway*


JAMES MONTANUS *University of Rochester*

THE ANALYTICAL ECONOMIST

Derivatives: Not the Real Thing

During the 1980s, junk bonds were perceived as either the snake oil that would destroy the stability of the financial world or the entrepreneurial grease that would lubricate wealth-building transactions old-line bankers were too stodgy to fund. Now so-called derivative securities seem to be playing the same role. They offer marvelous opportunities to make money, but just as junk bonds left many investors holding worthless paper, derivatives have their downside. Trading volume is well over $10 trillion—approaching the combined gross national product of the U.S., Japan and Europe—most of it unregulated because government agencies have yet to catch up. Increasingly, observers have begun to worry that a major misstep could vaporize financial markets.

Probably not, says Bhagwan Chowdhry of the University of California at Los Angeles. For one, the amount of money at risk is usually only a tiny fraction of the trading volume—as little as a few thousand dollars on a $100-million deal. For another, unlike real markets, derivative markets are zero-sum: for every big loser, there is also a big winner. Unless a player defaults (with debts exceeding assets), wealth can only be redistributed, not created or destroyed. Meanwhile, according to Yale University economist Stephen A. Ross, derivatives attract investors and so help money flow more smoothly through markets. This effect, he claims, is "an unadulterated good."

DANIEL NICHOLS *Liaison International*

EGGS OR HIGH FINANCE? The actual object underlying most derivatives is irrelevant to traders.

What is a derivative security anyway? The term does not come from the complex math that has made advanced degrees in physics or computer science so valuable on Wall Street recently. Instead it lumps together financial instruments whose common feature is that their value is ultimately derived from other securities, such as government bonds, stock in corporations or contracts to buy commodities such as gold, pork bellies or foreign currency.

The simplest derivatives are futures, contracts that set a price today but specify acceptance or delivery months hence. A sausage maker might buy hog futures to protect himself from the chance that prices may go up, whereas a meat packer would sell them to ensure against losses if the price goes down. A speculator might either buy or sell in the hope that a change will allow her to make money by reversing the transaction tomorrow.

Because every order to sell must be matched by an order to buy, derivative markets as a whole balance out to zero—as opposed to stock markets, where companies may issue shares regardless of whether there are buyers. As Chowdhry notes, the price of futures is constrained by the current cost of the underlying commodity. Otherwise, if the price of gold futures, say, rose above a certain point, speculators could profit by buying gold today and holding it for sale on the delivery date.

One notch up in complexity, options confer the right to buy or sell stock (or other securities) at a fixed price for some period. They are bets on the stock's price, and the cost of the option is the ante for getting into the game.

Simple derivatives such as futures and options have been around for centuries. In 1973 academic economists Fischer Black and Myron S. Scholes (now both on Wall Street) published their eponymous formula, which put the value of options on a firm mathematical foundation. Since then, economists and traders have developed far more complicated derivatives as well as analytical tools to set their prices.

Consider the interest-rate swap, as explained by Eduardo Schwartz, also at U.C.L.A. Company A wants to benefit from falling interest rates; B would like to protect itself against a possible rise. So A "lends" B $100 million at a fixed rate (say, 8 percent), and B "lends" A $100 million at a variable rate. Each month they balance accounts; if the variable rate is greater than 8 percent, A pays B the difference; if it is less, B pays A. Although the loan principal is recorded on the books of each company, Schwartz says, it is only "notional," an accounting fiction.

In the derivatives market, these swaps take on a value of their own. Number crunchers can look at today's interest rates, figure out how much income a swap will generate and for whom, and can then sell the swap for an appropriate price. To add another layer of complexity, companies may exchange the payments from debts denominated in different currencies, the income from U.S. treasury bills for dividends generated by a portfolio of Japanese stocks, or any other gizmo financial markets offer. Each combination allows the participants to trade a different set of potential risks and benefits, and the only limit to the complications is "the human mind," Ross quips.

What seems to be frightening regulators is evidence that the minds of some humans in the market may already have reached their limits. A Chilean government employee traded away nearly 0.5 percent of the nation's GNP by playing copper futures, and the German conglomerate Metallgesellschaft lost $1.3 billion betting on the future price of oil. Although the losers in these two cases were able to pay off the market winners, the sheer volume of obligations produces anxiety.

In theory, speculators can protect themselves against such risks by hedging—purchasing derivatives whose values rise or fall in opposition to one another, but in practice they may forgo the safety net for greater profits. If interest rates shifted sharply, for instance, the holder of one side of a swap might not be able to come up with the necessary payments, and then the fictional loans underlying the transaction would become an ugly reality. Although all but a tiny fraction of the interlocking debts would cancel out, even that could cause the holder of the other side of the swap to default on its obligations, triggering a general collapse. Because few rules govern the derivative markets, no one knows for sure how great, or small, the risk may be. —*Paul Wallich*

TECHNOLOGY AND BUSINESS

Bellcore on the Block

Second-largest U.S. industrial research center might be sold

Robert W. Lucky shakes off a shiver as he sits down to talk about the future of Bell Communications Research, the giant telecommunications laboratory where he presides over applied research. Earlier that November morning he had awakened to find his powerboat missing. High winds during the night had loosed it from its mooring on the Navesink River in New Jersey. Lucky found the wayward craft drifting downstream, intact but powerless, its batteries dead. He had no choice but to plunge into the chill waters and swim home with the vessel in tow.

LAUREN GREENFIELD *Sygma*

DISASTER RECOVERY is Bellcore's forte. But can it handle an industry shake-up?

The unpleasant adventure sounds a lot like Lucky's job description. Bellcore is struggling through a storm of competitive maneuvering as the Baby Bells that own it break their geographical bounds and grasp for national markets in wireless, video and data services. "The model that Bellcore was based on—seven owners with congruent interests—is certainly no longer the reality," Lucky observes. Increasing competition among the telephone companies "makes it almost untenable as time goes along. Each one wants proprietary solutions, unique advantages. And their biggest competitors are sitting across the [boardroom] table."

Many inside Bellcore and the larger research community believe the Bell companies—Pacific Bell, Southwestern Bell, U.S. West, Nynex, Bell Atlantic, BellSouth and Ameritech—are planning to cut the laboratory loose if a buyer can be found. George H. Heilmeier, the lab's president and CEO, calls such talk premature. The Bells "are still our owners," he says. "There has been no indication—or rather no decision has been made to sell Bellcore." But Lucky confirms that "it is being studied."

Some onlookers worry that "the great unmentionable," as one Bellcore spokesperson refers to the prospect, might jeopardize the integrity of the national telephone network. "Our industry has relied on Bellcore to perform certain functions," says Kathleen Levitz of the Federal Communications Commission. Bellcore sets standards to ensure that equipment from different vendors works together reliably, especially in an emergency, such as an earthquake. It also administers a database of toll-free services and the area code and prefix plan.

Yet in response to grumbling by non-Bell telephone companies, "there has already been an evolution from reliance on Bellcore for many of these functions to industry forums," Levitz reports. So long as any new owner has no major conflict of interest, it "would not need the FCC's approval to consummate a sale of Bellcore," she says.

New ownership would probably accelerate profound changes already under way at the lab. Since Heilmeier took over in 1991, he has tried to appease the Baby Bells' demands for more short-term results and has attempted to make the lab more self-sustaining. He cut Bellcore's staff by 28 percent, to 6,200. And he won permission from the directors to market Bellcore's services internationally to non-Bell companies. Such contracts now bring in about 16 percent of the lab's $1-billion annual revenues.

Most important, Heilmeier reversed Bellcore's research emphasis from hardware to software. That strategy partially circumvents one of Bellcore's greatest limitations: its legal prohibition, as a subsidiary of the Baby Bells, from designing or manufacturing tangible products. That restriction would vanish if the Bells sold their stakes in Bellcore. In the meantime, Lucky has supervised what he calls "a shift from the physical to the virtual." Software has grown to consume 70 percent of the lab's effort—and most of that is spent on production, rather than discovery and invention. Today "research is only 10 percent of Bellcore," Lucky says.

What good research remains may be jeopardized by the uncertainty over Bellcore's future. Until recently, the lab had a world-renowned group working on discrete math and theoretical computer science. The team was the pride of Bellcore. Such a resource "is very hard to cultivate but very easy to destroy," says Fan R. K. Chung, who assembled and led the crew but resigned this autumn. "It is widely known that we are in search of new owners," she states. "And there is no answer to the question of where we are going." Many of her dozen or so teammates have already jumped ship. "More than half have left or are leaving," she reports. "Of course, the best people leave first, because they have more options."

Bellcore's sale could free it of another onerous restriction, one imposed by the Baby Bells. By charter, Bellcore can work only on the regulated side of the telephone business. Unregulated areas such as personal communications services and video-on-demand—the growing markets in which new technology is most needed and valuable—have been ruled off limits.

Converting Bellcore from central research facility to contract lab could solve that problem. Some of Bellcore's directors already seem to view Bellcore much like a contractor. Ross Ireland, Pacific Bell's vice president for technology, points out that "there are a lot of good alternatives to Bellcore. Bell Labs, Bell Northern Research, Hewlett-Packard Labs and Xerox PARC are all doing similar work." John F. Gamba, a senior vice president at Bell Atlantic, says "they are becoming much more like just another player." And William B. Smith, chief information and technology officer for U.S. West Communications, notes that "as far as I'm concerned, they are in open competition for our business."

The flexibility and independence Bellcore might gain if it were set adrift would come at a steep price: the stability of its funding and thus the quality of its research. "Contract research poses a whole new minefield of difficulties," says Alan G. Chynoweth, who managed research at Bellcore from its conception in 1983 until his retirement in late 1992. "The opportunity is for Bellcore to become an industry-wide resource, not just for the regional Bells but for the nation, in making the information superhighways really work. Bellcore could have a tremendous role with its intimate knowledge of networks. But a contract basis is a much less secure base on which to support a quality research organization." Bellcore began imposing salary caps in November to cut costs.

Lucky worries about this trend as well. "We are a microcosm of what's happening throughout industry. I really am concerned about communications research in this country," he frowns. And shudders again. *—W. Wayt Gibbs*

The Rights Stuff

Buying and selling art in a digital world

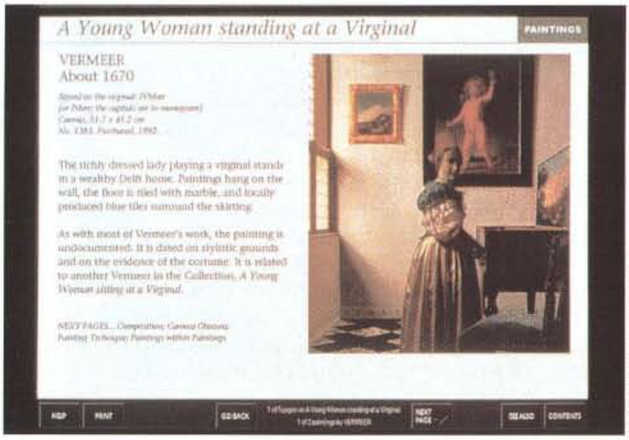
Museums and artists are not the only groups interested in getting art to the public. Motivated perhaps more by the possibility of enriching their coffers than of elevating the human spirit, software companies are embarking on a new kind of art acquisition, procuring digital rights to paintings, sculptures and other objects. Ownership of those rights allows a company to post works of art on electronic bulletin boards or to incorporate them in CD-ROMs and other multimedia products.

Museums have long granted publishers the right to reproduce images of their art holdings. But the electronic revolution is taking museums into uncharted legal and technological terrain. Curators worry about losing control of their collections: replicating and altering masterpieces are much easier when the images are in digital form. And museum directors, many of whom are unfamiliar with multimedia, have to decide how much to trust outside companies to disseminate the digitized versions of their holdings. By the end of 1995 there will be nearly 10 million CD-ROM-equipped computers in the U.S., twice as many as a year ago, so the stakes are high and rising.

Issues surrounding digital-art rights surfaced in the late 1980s, when museums and collections were contacted by a new player in the art world. Bill Gates, the founder of Microsoft Corporation, became interested in acquiring the rights to artworks that could form the basis for future software products. Because the market for those images was far from clear, Microsoft's board of directors declined to fund the venture; Gates therefore created an independent company, Interactive Home Systems (IHS), to carry out his goal.

Gates initially attempted to purchase exclusive digital rights; in other words, he asked museums to give up their freedom to sell digitized images to any other company or to develop commercial products of their own. Gates was quickly rebuffed; his commercial audacity ran contrary to both the content and tone of normal requests for permissions. Museum employees quietly mutter about the brusque and uninformed attitude of some IHS representatives. "You can call it cultural imperialism," laughs Benjamin H. Davis of the Massachusetts Institute of Technology. "But IHS was out to corner the market; that's what a company does."

IHS subsequently tamed its approach and has changed its name. In its new guise as Continuum Productions, it shops for nonexclusive rights and has sharpened its art-world connections by hiring as an adviser J. Carter Brown, former director of the National Gallery of Art in Washington, D.C. Continuum has struck deals with such collections as the Seattle Art Museum, the National Gallery in London and the Barnes Foundation. It currently has some 200,000 images in its collection, about 25,000 of which are fine art. Other holdings relate to science, music and history. Microsoft is counting on its clout in the software market to give it an edge in getting the images to the consumer.



A Young Woman standing at a Virginal

VERMEER
About 1670

The richly dressed lady playing a virginal stands in a wealthy Delft home. Paintings hang on the wall, the floor is tiled with marble, and locally produced blue tiles surround the skirting.

As with most of Vermeer's work, the painting is undocumented. It is dated on stylistic grounds and on the evidence of the costume. It is related to another Vermeer in the Collection, *A Young Woman sitting at a Virginal.*

NEXT PAGES... Composition; Camera Obscura; Painting Technique; Paintings within Paintings

DIGITIZED ARTWORKS challenge the way that museums present and distribute images of their holdings.

MICROSOFT CORPORATION

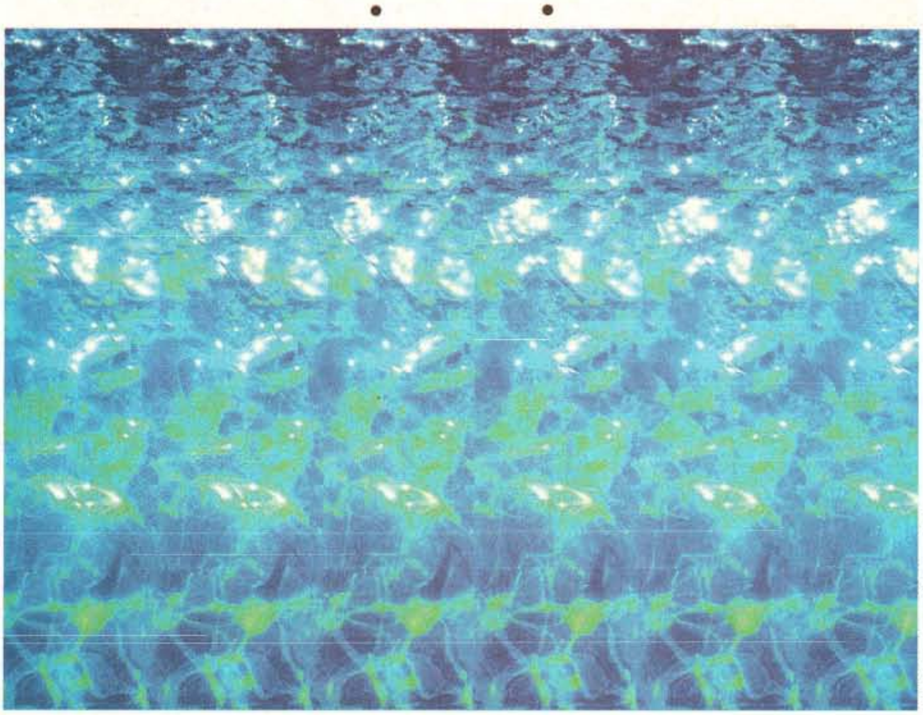
Who that consumer will be is still a bit murky. Stephen B. Davis, Continuum's director of strategic development, explains that the company's vast collection will form two kinds of products. The first will be essentially a multimedia stock agency; the second will be a series of databases that he hopes to make available to the public, either via CD-ROM or through an electronic network.

Not surprisingly, some of the first users of digitized art images are the museums themselves, eager to use the aura of high technology to lure new visitors. Three years ago the Seattle Art Museum unveiled ViewPoint, an interactive kiosk developed by Continuum. The touch-screen computer lets visitors browse through the holdings, much like using a computerized directory in a shopping mall. Britain's National Gallery recently installed a comparable setup, called Micro Gallery; the National Gallery in Washington, D.C., is developing a similar but more ambitious kiosk system. Both national galleries worked not with Continuum but with Cognitive Applications, a British software firm.

Microsoft has taken a leading role in getting digital art out of the museum and into the hands of the public. The company bought the rights to the British Micro Gallery and reengineered it into a CD-ROM. The resulting product, "Art Gallery," has been a commercial hit, selling 100,000 copies. Gates does not have a lock on the art-software industry, however. Digital Collections, for example, sells several art CD-ROMs, including one featuring works from the Frick Collection in New York City. Numerous digital-art encyclopedias, and even fine-art screen savers, are starting to appear in software stores.

A proliferation of other publishers is keeping the market in flux, and museums are, on the whole, carefully hedg-

Secrets in Stereogram



JUN OI courtesy of Cadence Books

They fit the profile of certain illicit drugs: ubiquitous, addictive, the cause of euphoria as well as irritability and lassitude. Like LSD, they are available as dots on paper. But while stereograms seem psychedelic in origin, they are actually the product of psychology, military research and art.

The popularity of three-dimensional viewing of two-dimensional images dates to the early 1800s, when the stereoscope was invented. It was not until the 1950s, however, that so-called random-dot stereograms, which resemble Jackson Pollack paintings, were created. At that time, according to the October 1994 issue of the *American Mathematical Monthly*, psychologist and engineer Bela Julesz looked through a stereoscope at two aerial photographs of camouflaged areas taken from slightly different angles. He noted that previously hidden tanks seemed to jump out. Julesz determined that depth perception did not take place in the eye but at a higher place in the brain.

In 1979 a former student of Julesz's, Christopher W. Tyler, created the single-picture stereogram, akin to the one shown here. Such images did not need the stereoscope; with a little training, eyes could find the hidden three-dimensional picture.

To see what all the hoopla is about, just cross your eyes until the two black dots above this image become four. Then, through luck or will, make the two central dots of your hallucination coalesce, until only three spots remain. Take an aspirin. Focus on the middle dot. When it is clear and unmoving, slowly bring your eyes down over the picture. You should see parts of the image bending concavely in the form of two marine mammals—unless you are part of the 2 percent of the population that is stereoblind. Of course, it may be simpler to just hang out of a helicopter and look for camouflaged army equipment. —*Marguerite Holloway*

ing their bets. "We've been approached by a lot of people. We're not doing anything that is either permanent or difficult to back away from," says Charles S. Moffett, director of the Phillips Collection in Washington, D.C. A few museums—especially small, technologically aware ones, such as the Michael C. Carlos Museum in Atlanta—are taking matters into their own hands by making their collections available on the Internet via the World Wide Web. Even the Smithsonian Institution is getting in the act, offering digital images through America Online.

Some critics worry that the limited resolution and poor color accuracy of computer monitors degrade the quality of the art. Most museums, however, seem firmly convinced that familiarity increases interest in the original work. A more serious concern involves maintaining control of images. One startling feature of Art Gallery—and of any unprotected CD-ROM—is that it allows users to copy the images off the disk and manipulate them on the computer. Drawing a mustache on the Mona Lisa has never been so easy. Defacing reproductions of great art is hardly a new game, but what is novel is the ability to create, save and erase the changes.

The legal departments of museums are still coming to grips with the implications of digitized artworks. They will have to determine which breaches of electronic rights they wish to pursue. "I don't know that there is a solution. Legally, you just go after the biggest offenders," says Alan B. Newman, executive director of imaging at the Art Institute of Chicago. Museums may try to attach copyright tags to images that they post through the Internet, as the Smithsonian does.

As Ben Davis points out, digitized art "is a medium you can literally do anything with: it's transmissible, it's alterable, you can make new art out of it." Already lawsuits are blooming over the appearance of pirated cartoon characters and scanned-in *Playboy* nudes on the Internet; manipulated fine-art images are also beginning to show up. The limited resolution of current CD-ROM or on-line images restricts their usefulness. But soon it will be easy to store and transmit publication-quality digitized artworks. Some computer-literate museum employees are starting to think about ways to encrypt such images so that only authorized users can look at them.

Last summer the Association of Art Museum Directors held a special meeting, "Art Museums on the Information Superhighway," to consider the philosophical implications of digital art and to sort out questions about image ownership in the electronic age. "The computer makes capitalism very transparent—it's all about property rights," Ben Davis reflects. "The problem is, museums don't see themselves in the art-information business." —*Corey S. Powell*

On the Road to Nowhere?

Management failures hold up the development of a clean car

Watchdog groups and conspiracy nuts alike wondered whether the Big Three automakers were merely deflecting political heat when they agreed in September 1993 to develop a superefficient "clean" car. The manufacturers have, after all, successfully defeated legislative attempts to reduce fuel consumption. A review published by the National Research Council in November seems to bear out the skeptics' fears: the corporate and federal collaboration, called the Partnership for a New Generation of Vehicles (PNGV), is earning mixed grades.

The NRC committee concludes that neither the government nor the companies have adequate management structures in place. That situation "may contribute to the apparent absence of specific program plans." In addition, because of the confidentiality of projects funded by industry, the NRC declared itself "unable to assess the suitability of the timing or adequacy of the industry funding." Although the PNGV partners have pledged to collaborate, the report goes on to charge that "antitrust laws and basic competitive and proprietary interests tend to limit the sharing of technologies and information between companies."

Nevertheless, not all the news is bad. On the technical side, the PNGV does appear to have considered how to reach its goals—although nobody will be bending metal or plastic to build clean cars in the near future. Plans call for the development by 2004 of a prototype suitable for mass production with three times the fuel efficiency of today's models. Although tripling mileage per gallon could be done with current technology, the PNGV version must perform as well as contemporary gas guzzlers and cost no more.

The official PNGV line is that no decisions have been made about design. But the prototype will probably be a "light hybrid," says Henry Kelly, an official in the Office of Science and Technology Policy. Electric motors, which can operate at high efficiency, are an obvious choice for a low-pollution vehicle. Batteries that could store energy for a full day's drive are, however, too heavy. Hybrid designs obviate this problem by using an auxiliary power source. This secondary unit produces electricity that drives the primary motor. The supplemental source might be a fuel cell, which produces electricity by combining hydrogen and oxygen. Or it might consist of a generator driven by a high-efficiency internal-combustion engine or gas turbine.

Hybrid cars already exist, although they are not close to achieving the PNGV's requirements on cost and performance. Reaching those targets will require at least one quantum jump in a critical technology, says Paul B. MacCready of AeroVironment in Monrovia, Calif. Nevertheless, he adds, plenty of promising approaches have not yet been investigated thoroughly. The PNGV target "requires a lot of government money, because until the leap happens, the process won't be market driven," MacCready argues.

The controller for the electric motor may represent the biggest challenge. According to Kelly, the only way it could operate well enough is for it to be a "smart" device that constantly assesses driving conditions and battery-charge levels. Capturing and storing the energy released in braking poses another challenge—although flywheels or high-power batteries might be up to the task. Energy losses must also be reduced: tires could be improved to cut resistance, and aerodynamic drag could be slashed by careful design. "All these things can be pushed further than they have been," MacCready notes.

Because PNGV's program calls for reducing the weight of vehicles by up to 40 percent, experts say high-tech materials will be crucial. Amory Lovins, research director at the Rocky Mountain Institute in Colorado, argues that ultralightweight compounds such as polymers could bring about reductions of more than 70 percent. If Lovins is right, future cars might weigh in at 400 kilograms, instead of the current 1,400 kilograms. Lovins calculates that the PNGV prototype could be between five and 20 times more efficient than vehicles now on the road.

Lovins also maintains it will be possible to hold down the price of such a vehicle. "The best way to make a car cheap is to use expensive materials," he proclaims. His logic is that only small amounts of such materials will be needed. Lifetime costs could be reduced if autos were equipped with more sophisticated systems to diagnose faults and engineered to ease service, Lovins speculates. The PNGV has not overtly embraced Lovins's ultralight gospel. But Lovins believes there is a "convergence" of opinion. "It's all right now for engineers to create original thoughts, and for the first time, military, aerospace and national lab composite experts are talking to the car guys," he says.

Even though the ideal car must be relatively cheap for consumers, deciding how to build it is expensive. The NRC report states that the PNGV needs further congressional support as well as a federally controlled line-item budget. So far the partnership has identified about $300 million a year of federal and industry research that could, in principle, lead it toward its goal, says Tim Adams of Chrysler Corporation. By next year the program could be coordinating research worth more than $500 million annually. But if the NRC is right, more money may do little: unless the bureaucrats get organized, the clean car will remain stalled. —*Tim Beardsley*

CHRYSLER CORPORATION

HYBRID RACING CAR recently designed by Chrysler Corporation uses a natural gas engine and an electromechanical battery, or flywheel. Passenger models using a hybrid approach are being studied by the Partnership for a New Generation of Vehicles.

Invasion of the Bean Counters

Physician profiles—the good, the bad and the unadjusted

How good is your doctor? You may be unable to answer that question, but many insurance companies and health maintenance organizations (HMOs) have no such lingering doubts. They are hiring and firing—the polite term is "deselecting"—doctors based on statistical analyses of their practice patterns. Patients may no longer be reimbursed for consulting deselected physicians.

Doctors in Texas, Tennessee, Washington, D.C., and elsewhere have challenged deselections in court. At issue are the criteria used to decide who stays and who goes. In the District of Columbia, Blue Cross and Blue Shield reportedly spared physicians with prestigious practices the detailed examinations that others underwent.

In another instance, Cigna initially dropped almost all its contracts with black doctors in Kansas City. The case is enlightening because no one contends the insurer used race as a criterion. Instead it cut off doctors who were too expensive; observers say the ranking just happened to hit one race particularly hard. Because medical needs vary sharply according to a patient's age and illness, physicians who treat an older or sicker population may cost a company more per patient than those who deal with the young and healthy.

Indeed, Susanne Salem-Schatz and her colleagues at the Harvard Community Health Plan studied their group's doctors and found that adjusting for case mix could completely reverse a physician's profile. They measured the percentage of patients that each doctor referred to a specialist—many profiles downgrade physicians who refer more patients because specialists are expensive. When the figures were adjusted for the age and sex of patients, a quarter of the doctors stood out as significantly more free-spending than average and a similar number as significantly less likely to refer.

Taking into account the severity of each patient's illness reduced the fraction of outliers almost by half. Even more important, some physicians who initially appeared to be parsimonious with referrals ended up at the spendthrift end of the new ranking.

Case-mix adjustment is crucial to figuring out how expensive doctors really are or how well they treat their patients—and it is often the only information available cheaply. But "there are a number of systematic flaws in insurance-based data," says Deborah W. Garnick of Brandeis University. The analyses are only as good as the diagnostic codes that physicians put on their forms.

Jonathan P. Weiner and his co-workers at Johns Hopkins University have developed a system that translates the billions of codes to 50 ambulatory care groups (ACGs) that portray the approximate severity of each patient's condition. A health plan can compare different doctors' performance on patients in the same ACG category and get a fairly good idea of their relative standing, Weiner says. Blue Cross/Blue Shield of North Carolina, for example, is using ACG profiles to determine bonuses for its HMO physicians.

In the past, doctors chided for spending too much money would simply contend that their patients were sicker, according to Don W. Bradley of North Carolina Blue Cross/Blue Shield. ACG adjustments have proved many of them right, he says, and for the remainder, the profiles have far more credibility. Meanwhile doctors who appeared exceptionally effective in unadjusted rankings can now be seen as the beneficiaries of healthy clients.

Dan L. Gunselman, who helped to put this case-mix adjustment system in place, comments that it still has its weaknesses, especially in accounting for catastrophic ailments. "If someone gets in an auto accident and needs $500,000 worth of treatments," he points out, no amount of adjustment will restore the primary physician's cost profile. Gunselman is looking for ways to make more sophisticated compensations.

There will always be some variation that cannot be accounted for. "No system will get the nuances," Weiner says. That is why he and others oppose using profiles alone to decide which doctors to hire and fire. Instead of making decisions based on ACGs or any single criterion, he contends, HMOs and insurers should use profiles as a quick way to find doctors whose records—good or bad—deserve further attention.

Salem-Schatz is of like mind. Physicians often do not know how their patterns of practice compare with those of their peers, she says, and medical managers' first use of profiles should be to provide that feedback. The North Carolina Blue Cross/Blue Shield HMO gives its doctors a report card every six months; it censures those who spend too little on their patients as well as those who appear to spend too much.

While health plans compile data on their doctors, consumer organizations have begun to gather information on the health plans. The National Committee for Quality Assurance has put together a report card of items that employers and patients can demand from insurance companies and HMOs to see how well they are doing. Deselection can go many ways. *—Paul Wallich*

Food for Thought

WE'RE FREE!
YES!
PUTRESCENT
WHOLE EGG SOLID

MICHAEL CRAWFORD

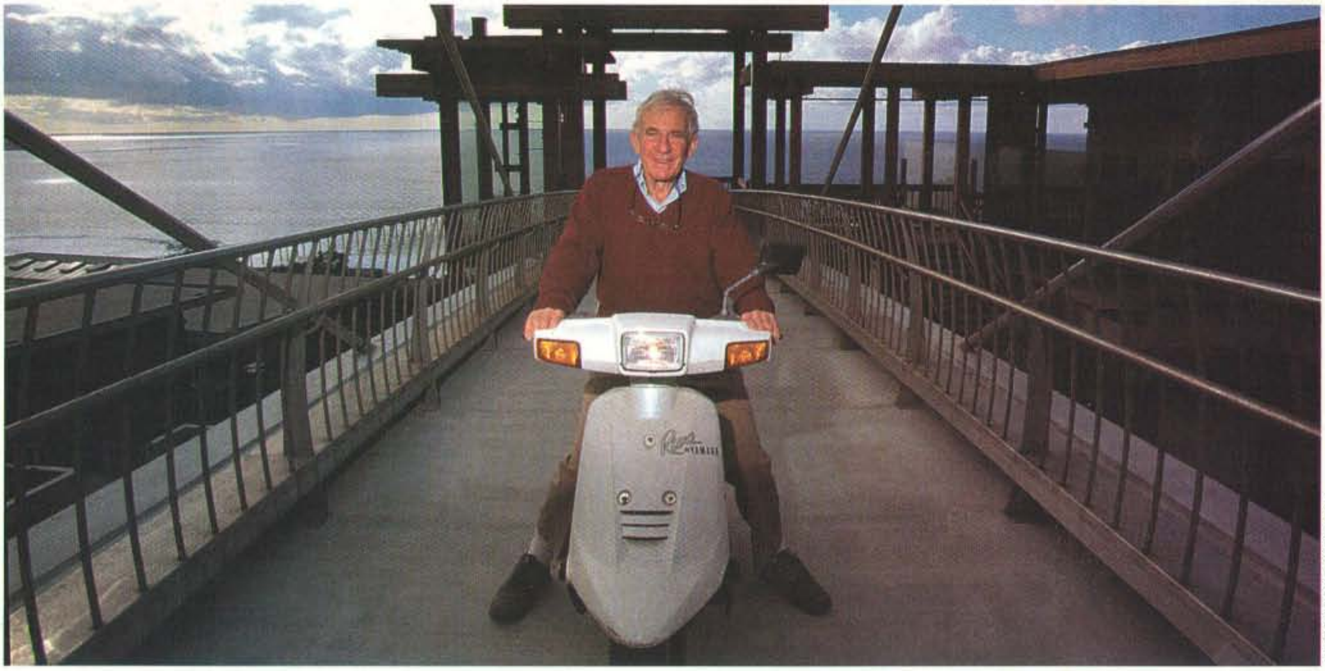
Alert the C.I.A., that is, the Culinary Institute of America. The Environmental Protection Agency has proposed removing several substances from its some 22,000-item-long list of registered pesticides. If the proposal is approved, companies or individuals wishing to use, distribute or sell any of the following compounds to kill unwelcome guests, such as vermin or weeds, will no longer be required to conduct the more than 75 toxicity tests often needed for each registered substance. They will no longer be obliged to endure the sometimes several-year process of having the pesticide approved and federally licensed. The pesticides under review include castor oil, cinnamon, cloves, corn oil, dried blood, garlic, mint, peppermint, putrescent whole egg solids, rosemary, sesame, soybean oil and white pepper. *—Marguerite Holloway*

PROFILE: WALTER H. MUNK

The Man Who Would Hear Ocean Temperatures

Walter H. Munk is ripping off his necktie. "I had to give a talk earlier," he explains almost apologetically. Southern California casual appears to rule in Munk's office, where the sound of the surf and the smell of the sea relax even the most anxiety-ridden visitor from the Northeast.

Despite the surrounding calm, the spry 77-year-old Munk charges ahead in his tasks, as he has for more than half a century at the Scripps Institution of Oceanography in La Jolla, Calif. Gracious and quick to smile, he leads me on a tour midway through our interview, eager to show off his institute and the cable-stayed bridge that connects it to the newer parts of Scripps. He explains how he and his wife, Judith, an architect by training, helped to design the center, whose buildings are carved into the uneven slopes of the coastline.


JAMES ARONOVSKY *Zuma*

OCEANOGRAPHER Walter H. Munk navigates by ship and scooter to study the earth and its waters.

Matters of the high seas, however, make Munk most comfortable. While I sip the coffee he has poured me, Munk checks his electronic mail. "We heard yesterday that the U.S. Navy is planning to close their listening station at Bermuda," he reports in his slightly Austrian accent. Dismantling the post—originally designed to locate Soviet submarines—would take place in less than two weeks, and Munk and his colleagues have sprung into action. "Our plan is to persuade the navy to give it to us," Munk comments. Scientists could then listen for undersea earthquakes and monitor the migration of marine mammals.

Munk's persistence is not surprising. Associates have described him—in a positive way—as a consummate salesman. His irresistible, infectious enthusiasm for what he does has won over many researchers and funding administrators. Indeed, Munk has been called one of the most influential oceanographers alive. "You say it in front of my wife tonight, and I know what she will say," Munk predicts. "Some four-letter word." (At dinner later, his wife resists, declaring she does not know me well enough.)

"What makes him a good scientist," remarks Carl Wunsch of the Massachusetts Institute of Technology and a longtime collaborator, "is his ability to see right through the math, to what it means physically." Munk's work has garnered him more than two dozen honors and awards, including the Vetlesen Prize, sometimes called the Nobel in earth science.

Nevertheless, Munk's stature received a bit of bruising recently. Environmental groups characterized his latest proposed experiment as deadly to marine mammals. To test climate models, Munk and his co-workers want to fire low-frequency sound waves off Kauai, Hawaii, and Point Sur, Calif. At a certain depth in the ocean, the temperature and pressure allow sound to travel thousands of kilometers without significant attenuation. Because sound moves faster in warm water than in cold water, changes in its average velocity can be measured over many years. The goal of the project, called acoustic thermometry of the ocean climate (ATOC), is to verify predictions by climate models that global warming is occurring.

Legal maneuvers and political action have already delayed the project by more than two years. Opponents argue that the rumbles could harm whales by disrupting their communication or by deafening and possibly killing them. "Certainly, whales can hear for several tens of kilometers, and it might interfere with their mating and feeding habits," Munk acknowledges. "It's a legitimate concern."

But one that has been blown out of proportion, the oceanographer insists. "It started out because there was a mistake made," Munk says. A postdoctoral student had the units wrong. "We would be transmitting 250 watts acoustic," Munk explains. "You don't physically damage at 250 watts, just as I don't physically damage you by talking to you." It would sound like a very loud orchestra a few meters away. "You wouldn't like it," he assures me, but the volume would do no harm. The student thought the level would be 250 million watts, which would be fatal to any life nearby. A story in the *Los Angeles Times* set off the reaction that threatened ATOC.

Munk's first oceanographic expedition, in 1940

COURTESY OF WALTER H. MUNK

The uproar caught the investigators off guard. "We've been working in the field for years without any problems," Munk points out. A dry run of ATOC in 1991 did not reveal any danger. Conducted off Heard Island near Antarctica in the southern Indian Ocean, the experiment blasted sound waves that were heard across the world, proving the feasibility of measuring ocean temperatures acoustically. Munk had arranged for marine biologists to monitor any effects on whales: "It was 1,000 times louder than what we want to do now, and we didn't cause any distress to the marine mammals." Munk also claims that other sources are far more disruptive. "We are about as loud as a tanker, and there are 1,000 tankers in the world. And tankers go 24 hours a day. As now proposed, we would be transmitting only 2 percent of the time, so we'd be very much less than a tanker."

Part of the trouble stems from language in the environmental impact statements, which declare that the experiment may "take" several hundred thousand mammals. In addition to death, the word meant any effect on behavior. "If you turn on your source and a whale changes its course by 10 degrees, you've taken him, by definition," Munk elaborates.

The controversy has abated, although at least one advocacy group remains, in Munk's words, "hostile." After obtaining the requisite permits, the ATOC workers hope to set sail this spring. "On the other hand, almost anyone can sue us," Munk observes. "You know, Scripps was concerned about the environment before the word 'environmentalist' had ever been used. To accuse the institution of being engaged in wholesale slaughter I think is terribly insulting."

Munk never anticipated that he would become an oceanographer. "I really grew up being interested only in skiing and tennis. Certainly not science," he states. His Viennese upbringing centered around finance. His grandfather was a banker who left enough money to provide for his children as well as a thriving branch in New York City. So at age 14 Munk was shipped to U.S. shores. "I was supposed to follow him," he laments. "My mother was kind enough to say that if I gave it a real try for a couple of years and didn't want it, I could do whatever I wanted. I didn't like it all. Gee, I never liked banks—they're boring." Munk chuckles. "The only time banks are willing to lend you money is when you don't need it."

Driven by ennui, Munk decided to get as far away from New York as possible. "I read the brochures and fell in love with those wonderful California names like Pasadena, San Marino. And the pictures looked very romantic." He ended up on the steps of the California Institute of Technology. "I was terribly naive," Munk reminisces. "I hadn't applied. I just showed up and knocked on the dean's door. I thought that was all it took." Perhaps amazed at the naiveté, the dean gave him an entrance examination, which Munk barely managed to pass.

Once enrolled, he studied applied physics, contemplating a career in geophysics. That notion quickly shifted. "I had a girlfriend whose grandparents were living in La Jolla, and she spent the summers there." Munk trailed her, taking a job at Scripps to pay for his living expenses. The woman dropped out of his life, but he liked Scripps so much that he returned to earn his doctorate under oceanographer Harald Sverdrup.

It was during World War II that Munk began a lifelong association with the navy. "I joined the army because I thought the end of the world was coming. Then the navy started some antisubmarine warfare," in which Roger Revelle, the late former director of Scripps, and Sverdrup were involved. They requested Munk be discharged from the army so that he could work alongside them. The switch was fortuitous. A few days later the Japanese attacked Pearl Harbor. "My unit had gone to New Guinea and was wiped out," Munk recalls. With Sverdrup, Munk predicted the occurrence of suitable waves that enabled Allied amphibious landings in northwestern Africa.

His military work constitutes only a small percentage of his contributions to earth science. "You'll see that I've been a dabbler," Munk remarks. "I do something for 10 years, then I do something else." With Scripps geophysicist Gordon J. MacDonald, he explained in the 1950s why the earth's axis wobbles and its spin varies slightly. In the 1960s he showed that storms near Antarctica give rise to the long, regular train of swells that rolls into southern California during the summer. In the 1970s he worked with Wunsch to develop ocean acoustic tomography. The technique, which relies on sound waves to create three-dimensional maps of ocean temperature and currents, led Munk directly to his present work on ocean climate.

"The inevitable outcome is that I don't do anything very well, because I don't stick with it long enough," Munk chides himself. "I'm not much of a scholar. I don't like to read. I like to work in a field that has nothing published, where you have to figure it out for yourself."

After dinner, Judith Munk leads the way to the deck to show me what they have chiseled into their backyard: an elegant amphitheater, large enough to accommodate 100 guests. Having been stricken with polio, she relies on a wheelchair for mobility. "We live very near to Jonas Salk," Walter mentions, "and we often accuse him if he hadn't been so damn lazy, if he had gotten his thing out a couple of years sooner, Judy wouldn't have come down with it." Munk laughs and throws up his hands: "He pleads guilty."

Although not a scientist, Judith has been instrumental in Walter's career—from taking the 4 A.M. ocean-swell watch in the Samoa Islands to influencing his thinking. "She has tremendously good common sense," Munk says. "She tells me when I do something stupid." Neither of the couple's two daughters is a scientist, although Walter likes to point out that one is married to a chemist.

On the deck, Judith encourages me to remove the drop cloth draped over the telescope that points out to sea. Only the light from a distant helicopter pierces the dark Pacific sky. By day the view of the ocean must be spectacular. "I love going to sea," Walter Munk muses. "It's a wonderful job." *—Philip Yam*

SCIENTIFIC
AMERICAN

Ensuring the Longevity of Digital Documents

The digital medium is replacing paper in a dramatic record-keeping revolution. But such documents may be lost unless we act now

by Jeff Rothenberg

The year is 2045, and my grandchildren (as yet unborn) are exploring the attic of my house (as yet unbought). They find a letter dated 1995 and a CD-ROM. The letter says the disk contains a document that provides the key to obtaining my fortune (as yet unearned). My grandchildren are understandably excited, but they have never before seen a CD—except in old movies. Even if they can find a suitable disk drive, how will they run the software necessary to interpret what is on the disk? How can they read my obsolete digital document?

This imaginary scenario reveals some fundamental problems with digital documents. Without the explanatory letter, my grandchildren would have no reason to think the disk in my attic was worth deciphering. The letter possesses the enviable quality of being readable with no machinery, tools or special knowledge beyond that of English. Because digital information can be copied and recopied perfectly, it is often extolled for its supposed longevity. The truth, however, is that because of changing hardware and software, only the letter will be immediately intelligible 50 years from now.

Information technology is revolutionizing our concept of record keeping in an upheaval as great as the introduction of printing, if not of writing itself. The current generation of digital records has unique historical significance. Yet these documents are far more fragile than paper, placing the chronicle of our entire period in jeopardy.

My concern is not unjustified. There have already been several potential disasters. A 1990 House of Representatives report describes the narrow escape of the 1960 U.S. Census data. The tabulations were originally stored on tapes that became obsolete faster than expected as revised recording formats supplanted existing ones (although most of the information was successfully transferred to newer media). The report notes other close calls as well, involving tapes of the Department of Health and Human Services; files from the National Commission on Marijuana and Drug Abuse, the Public Land Law Review Commission and other agencies; the Combat Area Casualty file containing P.O.W. and M.I.A. records for the Vietnam War; and herbicide information needed to analyze the impact of Agent Orange. Scientific data are in similar jeopardy, as irreplaceable records of numerous experiments conducted by the National Aeronautics and Space Administration and other organizations age into oblivion.

So far the undisputed losses are few. But the significance of many digital documents—those we consider too unimportant to archive—may become apparent only long after they become unreadable. Unfortunately, many of the traditional methods developed for archiving printed matter are not applicable to electronic files. The content and historical value of thousands of records, databases and personal documents may be irretrievably lost to future generations if we do not take steps to preserve them now.

From Here to Eternity

Although digital information is theoretically invulnerable to the ravages of time, the physical media on which it is stored are far from eternal. If the optical CD in my attic were a magnetic disk, attempting to read it would probably be futile. Stray magnetic fields, oxidation and material decay can easily erase such disks. The contents of most digital media evaporate long before words written on high-quality paper. They often become unusably obsolete even sooner, as media are superseded by new, incompatible formats—how many readers remember eight-inch floppy disks? It is only slightly facetious to say that digital information lasts forever—or five years, whichever comes first.

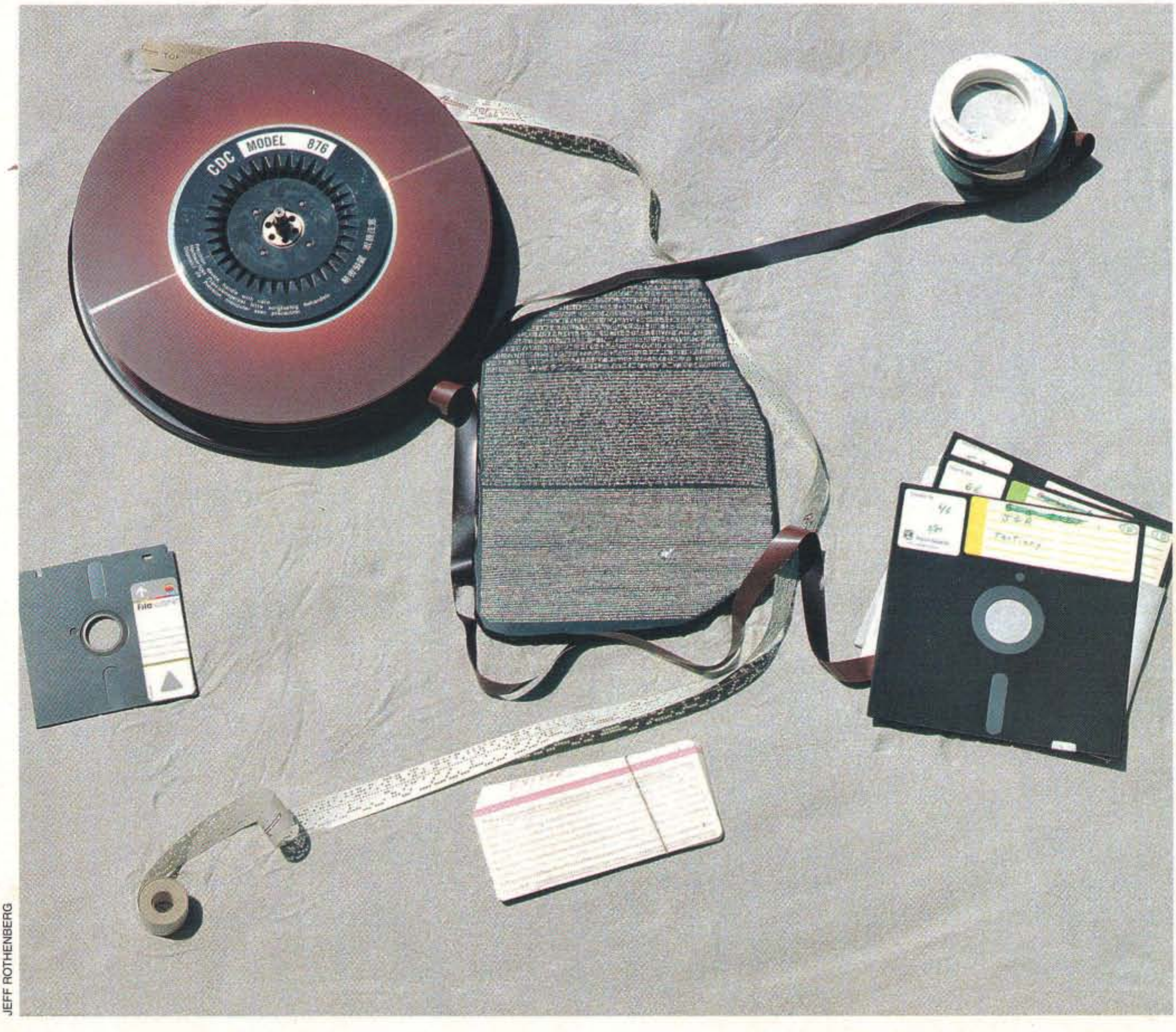
Yet neither the physical fragility of digital media nor their lemminglike tendency toward obsolescence constitutes the worst of my grandchildren's problems. My progeny must not only extract the content of the disk but must also interpret it correctly. To understand their predicament, we need to examine the nature of digital storage. Digital infor-

JEFF ROTHENBERG is a senior computer scientist in the social policy department of the RAND Corporation in Santa Monica, Calif. He received a master's degree in computer science from the University of Wisconsin in 1969 and then spent the next four years working toward a doctorate in artificial intelligence. His research has included work in modeling theory, investigations into the effects of information technology on humanities research, and numerous studies involving information technology policy issues. His passions include classical music, traveling, photography and sailing.

mation can be saved on any medium that is able to represent the binary digits ("bits") 0 and 1. We will call an intended, meaningful sequence of bits, with no intervening spaces, punctuation or formatting, a bit stream.

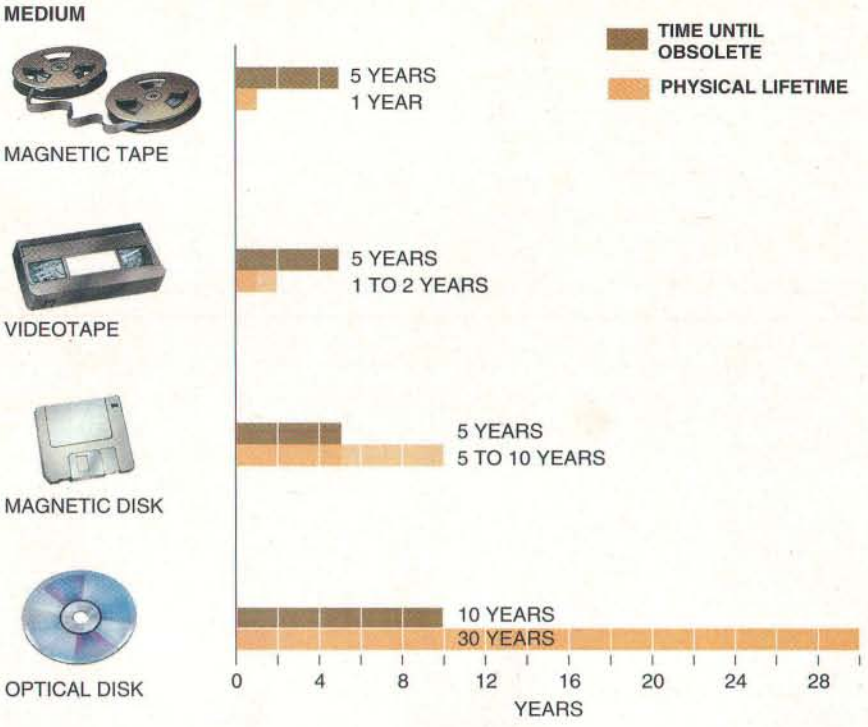
Retrieving a bit stream requires a hardware device, such as a disk drive, and special circuitry for reading the physical representation of the bits from the medium. Accessing the device from a given computer also requires a "driver" program. After the bit stream is retrieved, it must still be interpreted. This task is not straightforward, because a given bit stream can represent almost anything—from a sequence of integers to an array of dots in a pointillist-style image.

Furthermore, interpreting a bit stream depends on understanding its implicit structure, which cannot explicitly be represented in the stream. A bit stream that represents a sequence of alphabetic characters may consist of fixed-length chunks ("bytes"), each representing a code for a single character. For instance, in one current scheme, the eight bits 01110001 stand for the letter q. To extract the bytes from the bit stream, thereby "parsing" the stream into its



JEFF ROTHENBERG

OBSOLESCENCE plagues digital media. Those shown have already failed to remain readable for one hundredth the time that the Rosetta Stone has. The classical Greek script in the stone, which was found in 1799 in Egypt by a French military demolition squad, made hieroglyphics and demotic Egyptian comprehensible. Besides being legible after 22 centuries, the Rosetta Stone (a replica here) owes its preservation to the visual impact of its content—an attribute absent in digital media.



JANA BRENNING

EXPECTED LIFETIMES of common digital storage media are estimated conservatively to guarantee that none of the data are lost. (Analog tapes, such as those used for audio recordings, remain playable for many years because they record more robust signals that degrade more gradually.) The estimated time to obsolescence for each medium refers to a particular recording format.

components, we must know the length of a byte.

One way to convey the length is to encode a "key" at the beginning of the bit stream. But this key must itself be represented by a byte of some length. A reader therefore needs another key to understand the first one. Computer scientists call the solution to such a recursive problem a "bootstrap" (from the fanciful image of pulling oneself up by the bootstraps). In this case, a bootstrap must provide some context, which humans can read, that explains how to interpret the digital storage medium. For my grandchildren, the letter accompanying the disk serves this role.

After a bit stream is correctly parsed, we face another recursive problem. A byte can represent a number or an alphabetic character according to a code. To interpret such bytes, therefore, we need to know their coding scheme. But if we try to identify this scheme by inserting a code identifier in the bit stream itself, we will need another code identifier to interpret the first one. Again, human-readable context must serve as a bootstrap.

Even more problematic, bit streams may also contain complex cross-referencing information. The stream is often stored as a collection, or file, of bits that contains logically related but physically separate elements. These elements are linked to one another by internal references, which consist of pointers to other elements or of patterns to be matched. (Printed documents exhibit similar schemes, in which page numbers serve as pointers.)

Interpreting a Bit Stream

Suppose my grandchildren manage to read the bit stream from the CD-ROM. Only then will they face their real challenge: interpreting the information embedded in the bit stream. Most files contain information that is meaningful solely to the software that created them. Word-processing files embed format instructions describing typography, layout and structure (titles, chapters and so on). Spreadsheet files embed formulas relating their cells. So-called hypermedia files contain information identifying and linking text, graphics, sound and temporal data.

For convenience, we call such embedded information—and all other aspects of a bit stream's representation, including byte length, character code and structure—the encoding of a document file. These files are essentially programs: instructions and data that can be interpreted only by appropriate software. A file is not a document in its own right—it merely describes a document that comes into existence when the file is interpreted by the program that produced it. Without this program (or equivalent software), the document is a cryptic hostage of its own encoding.

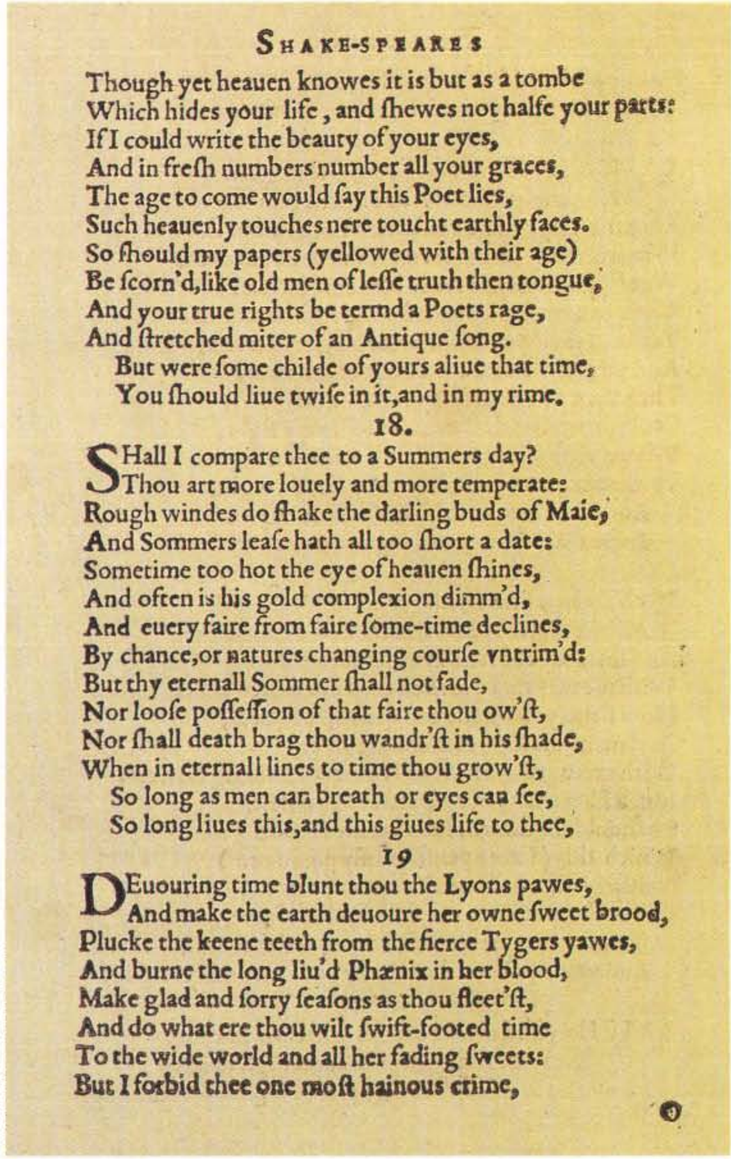
Trial-and-error might decode the intended text if the document is a simple sequence of characters. But if it is complex, such a brute-force approach is unlikely to succeed. The meaning of a file is not inherent in the bits themselves, any more than the meaning of this sentence is inherent in its words. To understand any document, we must know what its content signifies in the language of its intended reader. Unfortunately, the intended reader of a document file is a program. Documents such as multimedia presentations are impossible to read without appropriate software: unlike printed words, they cannot just be "held up to the light."

Is it necessary to run the specific program that created a document? In some cases, similar software may at least partially be able to interpret the file. Still, it is naive to think that the encoding of any document—however natural it seems to us—will remain readable by future software for very long. Information technology continually creates new schemes, which often abandon their predecessors instead of subsuming them.

A good example of this phenomenon occurs in word processing. Most such programs allow writers to save their work as simple text, using the current seven-bit American Standard Code for Information Interchange (or ASCII). Such text would be relatively easy to decode in the future if seven-bit ASCII remains the text standard of choice. Yet ASCII is by no means the only popular text standard, and there are proposals to extend it to a 16-bit code (to encompass non-English alphabets). Future readers may therefore not be able to guess the correct text standard. To complicate matters, authors rarely save their work as pure text. As Avra Michelson, then at the National Archives, and I pointed out in 1992, authors often format digital documents quite early in the writing process and add figures and footnotes to provide more readable and complete drafts.

If "reading" a document means simply extracting its content—without its original form—then we may not need to run the original software. But content can be lost in subtle ways. Translating word-processing formats, for instance, often displaces or eliminates headings, captions or footnotes. Is this merely a loss of structure, or does it impinge on content? If we transform a spreadsheet into a table, deleting the formulas that relate the table's entries to one another, have we affected content? Suppose the CD in my attic contains a treasure map depicted by the visual patterns of word and line spacings in my original digital version of this article. Because these patterns are artifacts of the formatting algorithms of my software, they will be visible only when the digital version is viewed using my original program. If we need to view a complex document as its author viewed it, we have little choice but to run the software that generated it.

What chance will my grandchildren have of finding that software 50 years from now? If I include a copy of the program on the CD, they must still find the operating system software that allows the program to run on some computer. Storing a copy of the operating system on the CD may help, but the computer hardware required to run it will have long since become obsolete. What kind of digital Rosetta Stone can I leave to provide the key to understanding the contents of my disk?



SHAKE-SPEARES

Though yet heauen knowes it is but as a tombe
Which hides your life, and ſhewes not halfe your parts:
If I could write the beauty of your eyes,
And in freſh numbers number all your graces,
The age to come would ſay this Poet lies,
Such heauenly touches nere toucht earthly faces.
So ſhould my papers (yellowed with their age)
Be ſcorn'd,like old men of leſſe truth then tongue,
And your true rights be termd a Poets rage,
And ſtretched miter of an Antique ſong.
But were ſome childe of yours aliue that time,
You ſhould liue twiſe in it,and in my rime.

18.

SHall I compare thee to a Summers day?
Thou art more louely and more temperate:
Rough windes do ſhake the darling buds of Maie,
And Sommers leaſe hath all too ſhort a date:
Sometime too hot the eye of heauen ſhines,
And often is his gold complexion dimm'd,
And euery faire from faire ſome-time declines,
By chance,or natures changing courſe vntrim'd:
But thy eternall Sommer ſhall not fade,
Nor looſe poſſeſſion of that faire thou ow'ſt,
Nor ſhall death brag thou wandr'ſt in his ſhade,
When in eternall lines to time thou grow'ſt,
So long as men can breath or eyes can ſee,
So long liues this,and this giues life to thee,

19

DEuouring time blunt thou the Lyons pawes,
And make the earth deuoure her owne ſweet brood,
Plucke the keene teeth from the fierce Tygers yawes,
And burne the long liu'd Phænix in her blood,
Make glad and ſorry ſeaſons as thou fleet'ſt,
And do what ere thou wilt ſwift-footed time
To the wide world and all her fading ſweets:
But I forbid thee one moſt hainous crime,

FOLGER SHAKESPEARE LIBRARY

SHAKESPEARE'S first printed edition of sonnet 18 (1609) exemplifies the longevity of the printed page: the words are legible after almost four centuries (the final couplet is especially relevant to preserving documents). But digital media can become unreadable within a decade.

Migrating Bits

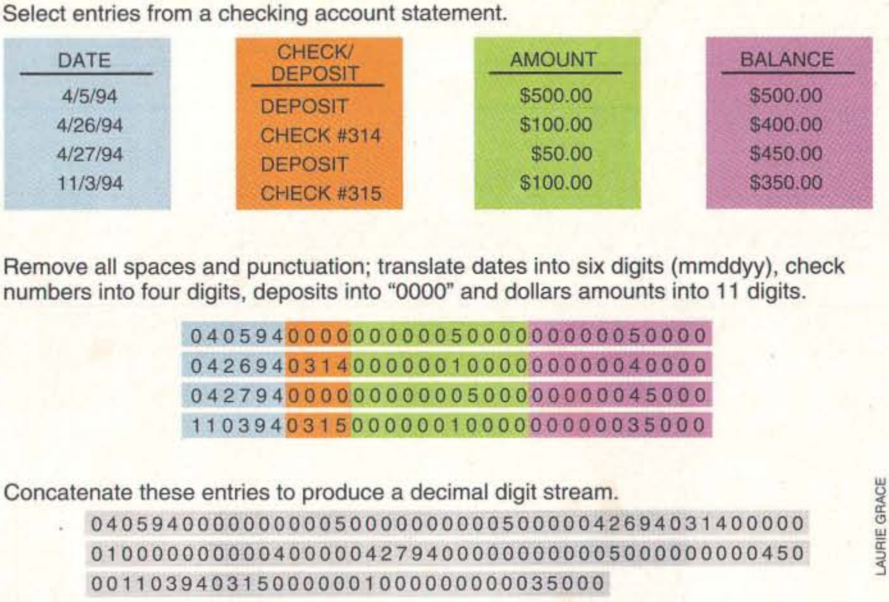
To prevent digital documents from being lost, we must first preserve their bit streams. That means copying the bits onto new forms of media to ensure their accessibility. The approach is analogous to preserving text, which must be transcribed periodically. Both activities require ongoing effort: future access depends on an unbroken chain of such migrations frequent enough to prevent media from becoming physically unreadable or obsolete before they are copied. A single break in this chain renders digital information inaccessible, short of heroic effort. Given the current lack of permanence of media and the rate at which their forms evolve, migration may need to be as frequent as once every few years. Conservative estimates suggest that data on digital magnetic tape should be copied once a year to guarantee that none of the information is lost. (Analog tapes may remain playable for many years because they record more robust signals that degrade more gradually.)

In the long run, we might be able to develop long-lived storage media, which would make migration less urgent. At the moment, media with increased longevity are not on the horizon. Nevertheless, the cost of migration may eventually force the development of such products, overriding our appetite for improved performance.

An ancient text can be preserved either by translating it into a modern language or by copying it in its original dialect. Translation is attractive because it avoids the need to retain knowledge of the text's original language, yet few scholars would praise their predecessors for taking this approach. Not only does translation lose information, it also makes it impossible to determine what information has been lost, because the original is discarded. (In extreme cases, translation can completely undermine content: imagine blindly translating both languages in a bilingual dictionary into a third language.) Conversely, copying text in its original language (saving the bit stream) guarantees that nothing will be lost. Of course, this approach assumes that knowledge of the original language is retained.

Archivists have identified two analogous strategies for preserving digital documents. The first is to translate them into standard forms that are independent of any computer system. The second approach is to extend the longevity of computer systems and their original software to keep documents readable. Unfortunately, both strategies have serious shortcomings.

On the surface, it appears preferable to translate digital documents into standard forms that would remain readable in the future, obviating the need to run obsolete software. Proponents of this approach offer the relational database (introduced in the 1970s by E. F. Codd, now at Codd & Date, Inc., in San Jose, Calif.) as a paradigmatic example. Such a database consists of tables representing relations among entities. A database of employees might contain a table having columns for employee names and their departments. A second table in the database might have department names in its first column, department sizes in its second column and the name of the department head in a third. The relational model defines a set of formal operations that make it possible to combine the relations in these tables—for example, to find the name of an employee's department head.



Select entries from a checking account statement.

DATE	CHECK/ DEPOSIT	AMOUNT	BALANCE
4/5/94	DEPOSIT	$500.00	$500.00
4/26/94	CHECK #314	$100.00	$400.00
4/27/94	DEPOSIT	$50.00	$450.00
11/3/94	CHECK #315	$100.00	$350.00

Remove all spaces and punctuation; translate dates into six digits (mmddyy), check numbers into four digits, deposits into "0000" and dollars amounts into 11 digits.

040594	0000	00000050000	00000050000
042694	0314	00000010000	00000040000
042794	0000	00000005000	00000045000
110394	0315	00000010000	00000035000

Concatenate these entries to produce a decimal digit stream.

```
040594000000000050000000000050000042694031400000
010000000000400000427940000000000005000000000450
00110394031500000010000000000035000
```

LAURIE GRACE

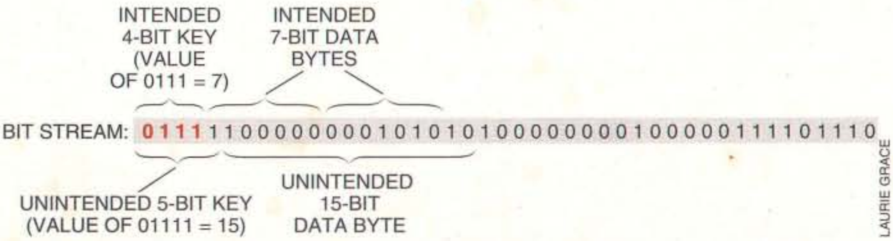
UNDERSTANDING A BIT STREAM demands knowledge of the format used to create the stream. If all the numbers in a monthly checking account statement were strung together—with nothing to distinguish check numbers, dates and dollar amounts—the resulting sequence of digits would be impossible to understand.

Because all relational database systems implement this same underlying model, any such database can in principle be translated into a standard tabular form acceptable to any other system. Files represented this way could be copied to new media as necessary, and the standard would ensure readability forever.

Flaws of Translation

Regrettably, this approach is flawed in two fundamental ways. First, relational databases are less standardized than they appear. Commercial relational database systems distinguish themselves from one another by offering features that extend the relational model in nonstandard ways. Moreover, the limitations of such databases are already leading to the adoption of new models. The tables in a relational database cannot transparently show structure. That is, the database could not immediately make it clear that a corporation consisted of one headquarters, five national offices, 25 divisions and 100 departments. Various object-oriented database models (which can represent structure directly) are evolving to satisfy this need. Such rapid evolution is neither accidental nor undesirable. It is the hallmark of information technology.

Furthermore, far from being a representative example, relational databases are practically unique. No other type of digital document has nearly so formal a basis for standardization. Word processors, graphics programs, spreadsheets and hypermedia programs each create far more varied documents. The incompatibility of word-processing files exemplifies this problem. It did not arise simply because companies were trying to distinguish their products in the marketplace. Rather it is a direct outgrowth of the technology's tendency to adapt itself to the emerging needs of users.


INTENDED 4-BIT KEY (VALUE OF 0111 = 7)
INTENDED 7-BIT DATA BYTES
BIT STREAM: 01111100000000101010100000000100000111101110
UNINTENDED 5-BIT KEY (VALUE OF 01111 = 15)
UNINTENDED 15-BIT DATA BYTE

LAURIE GRACE

CODE KEY may be used to indicate how a bit stream is organized. Here the first four bits stand for the integer 7, meaning that the remaining bytes are each seven bits long. Yet there is no way to tell the length of the code key from the bit stream itself. If we were to read the first five bits as the code key, we would erroneously conclude that the remaining bytes were 15 bits long.

As yet, no common application is ready to be standardized. We do not have an accepted, formal understanding of the ways that humans manipulate information. It is therefore premature to attempt to enumerate the most important kinds of digital applications, let alone to circumscribe their capabilities through standards. Forcing users to accept the limitations imposed by such standards or restricting all digital documents to contain nothing but text as a lowest common denominator would be futile. The information revolution derives its momentum precisely from the attraction of new capabilities. Defining long-term standards for digital documents may become feasible when information science rests on a more formal foundation, but such standards do not yet offer a solution.

Translating a document into successive short-term standards offers false hope. Successive translation avoids the need for ultimate standards, but each translation introduces new losses. Would a modern version of Homer's *Iliad* have the same literary impact if it had been translated through a series of intermediate languages rather than from the earliest surviving texts in ancient Greek? In theory, translating a document through a sequence of standards should enable scholars to reconstruct the original document. Yet that requires each translation to be reversible without loss, which is rarely the case.

Finally, translation suffers from a fatal flaw. Unlike English and ancient Greek, whose expressive power and semantics are roughly equivalent, digital documents are evolving so rapidly that shifts in the forms of documents must inevitably arise. New forms do not necessarily subsume their predecessors or provide compatibility with previous formats. Old documents cannot always be translated into unprecedented forms in meaningful ways, and translating a current file back into a previous form is frequently impossible. For example, many older, hierarchical databases were completely redesigned to fit the relational model, just as relational databases are now being restructured to fit emerging object-oriented models. Shifts of this kind make it difficult or meaningless to translate old documents into new standard forms.

The alternative to translating a digital document is to view it by using the program that produced it. In theory, we might not actually have to run this software. If we could describe its behavior in a way that does not depend on any particular computer system, future generations could re-create the behavior of the software and thereby read the document. But information science cannot yet describe the behavior of software in sufficient depth for this approach to work, nor is it likely to be able to do so in the near future. To replicate the behavior of a program, there is currently little choice but to run it.

For this reason, we must save the programs that generate our digital documents, as well as all the system software required to run those programs. Although this task is monumental, it is theoretically feasible. Authors often include an appropriate application program and operating system to help recipients read a digital document. Some applications and system software may remain ubiquitous, so that authors would need only to refer readers to those programs. Free, public-domain software is already widely available on the Internet. Moreover, when proprietary programs become obsolete, their copyright restrictions may expire, making them available to future users.

How can we provide the hardware to run antiquated systems and application software? A number of specialized museums and "retro-computing" clubs are attempting to maintain computers in working condition after they become obsolete. Despite a certain undeniable charm born of its technological bravado, this method is ultimately futile. The cost of repairing or replacing worn out components (and retaining the expertise to do so) must inevitably outweigh the demand for any outmoded computer.

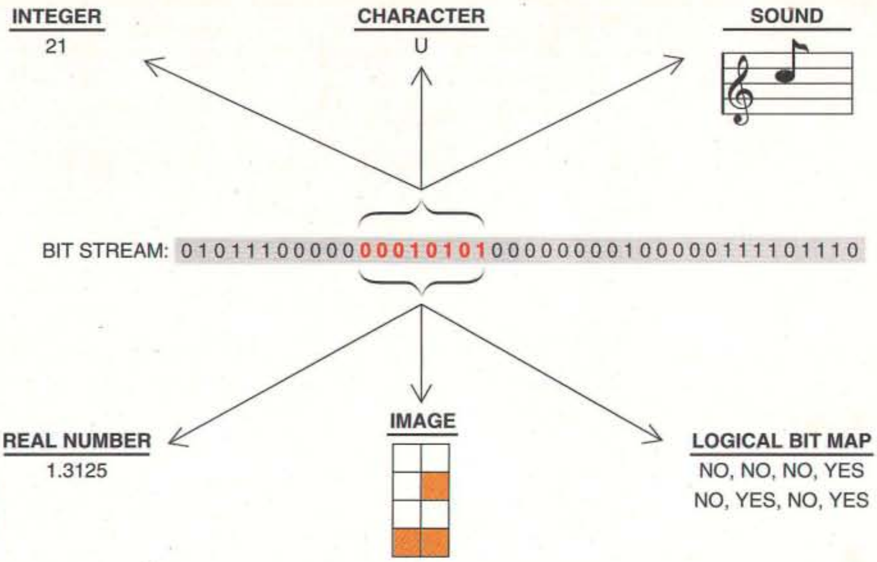
Fortunately, software engineers can write programs called emulators, which mimic the behavior of hardware. Assuming that computers will become far more powerful than they are today, they should be able to emulate obsolete systems on demand. The main drawback of emulation is that it requires detailed specifications for the outdated hardware. To be readable for posterity, these specifications must be saved in a digital form independent of any particular software, to prevent having to emulate one system to read the specifications needed to emulate another.

Saving Bits of History

If digital documents and their programs are to be saved, their migration must not modify their bit streams, because programs and their files can be corrupted by the slightest change. If such changes are unavoidable, they must be reversible without loss. Moreover, one must record enough detail about each transformation to allow reconstruction of the original encoding of the bit stream. Although bit streams can be designed to be immune to any expected change, future migration may introduce unexpected alterations. For example, aggressive data compression may convert a bit stream into an approximation of itself, precluding a precise reconstruction of the original. Similarly, encryption makes it impossible to recover an original bit stream without the decryption key.

Ideally, bit streams should be sealed in virtual envelopes: the contents would be preserved verbatim, and contextual information associated with each envelope would describe those contents and their transformation history. This information must itself be stored digitally (to ensure its survival), but it must be encoded in a form that humans can read more simply than they can the bit stream itself, so that it can serve as a bootstrap. Therefore, we must adopt bootstrap standards for encoding contextual information; a simple, text-only standard should suffice. Whenever a bit stream is copied to new media, its associated context may be translated into an updated bootstrap standard. (Irreversible translation would be acceptable here, because only the semantic content of the original context need be retained.) These standards can also be used to encode the hardware specifications needed to construct emulators.

Where does this leave my grandchildren? If they are fortunate, their CD may still be readable by some existing disk drive, or they may be resourceful enough to construct one, using information in my letter. If I include all the relevant software on the disk, along with complete, easily decoded specifications for the required hardware, they should be able to generate an emulator to run the original software that will display my document. I wish them luck.



LAURIE GRACE

INTERPRETING A BIT STREAM correctly is impossible without contextual information. This eight-bit sequence can be interpreted in at least six different ways.

FURTHER READING

Text and Technology: Reading and Writing in the Electronic Age. Jay David Bolter in *Library Resources and Technical Services,* Vol. 31, No. 1, pages 12–23; January–March 1987.

Taking a Byte out of History: The Archival Preservation of Federal Computer Records. Report 101-978 of the U.S. House of Representatives Committee on Government Operations, November 6, 1990.

Archival Management of Electronic Records. Edited by David Bearman. Archives and Museum Informatics, Pittsburgh, 1991.

Understanding Electronic Incunabula: A Framework for Research on Electronic Records. Margaret Hedstrom in *American Archivist,* Vol. 54, No. 3, pages 334–354; Summer 1991.

Archival Theory and Information Technologies: The Impact of Information Technologies on Archival Principles and Practices. Charles M. Dollar. Edited by Oddo Bucci. Information and Documentation Series No. 1, University of Macerata, Italy, 1992.

Scholarly Communication and Information Technology: Exploring the Impact of Changes in the Research Process on Archives. Avra Michelson and Jeff Rothenberg in *American Archivist,* Vol. 55, No. 2, pages 236–315; Spring 1992.

The Prion Diseases

Prions, once dismissed as an impossibility, have now gained wide recognition as extraordinary agents that cause a number of infectious, genetic and spontaneous disorders

by Stanley B. Prusiner

Fifteen years ago I evoked a good deal of skepticism when I proposed that the infectious agents causing certain degenerative disorders of the central nervous system in animals and, more rarely, in humans might consist of protein and nothing else. At the time, the notion was heretical. Dogma held that the conveyers of transmissible diseases required genetic material, composed of nucleic acid (DNA or RNA), in order to establish an infection in a host. Even viruses, among the simplest microbes, rely on such material to direct synthesis of the proteins needed for survival and replication.

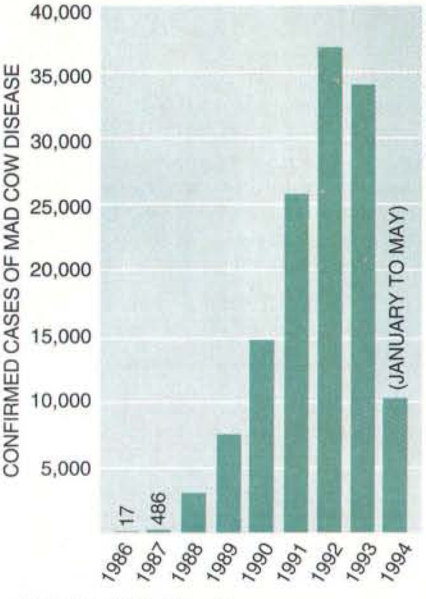
Later, many scientists were similarly dubious when my colleagues and I suggested that these "proteinaceous infectious particles"—or "prions," as I called the disease-causing agents—could underlie inherited, as well as communicable, diseases. Such dual behavior was then unknown to medical science. And we met resistance again when we concluded that prions (pronounced "pree-ons") multiply in an incredible way; they convert normal protein molecules into dangerous ones simply by inducing the benign molecules to change their shape.

Today, however, a wealth of experimental and clinical data has made a convincing case that we are correct on all three counts. Prions are indeed responsible for transmissible and inherited disorders of protein conformation. They can also cause sporadic disease, in which neither transmission between individuals nor inheritance is evident. Moreover, there are hints that the prions causing the diseases explored thus far may not be the only ones. Prions made of rather different proteins may contribute to other neurodegenerative diseases that are quite prevalent in humans. They might even participate in illnesses that attack muscles.

The known prion diseases, all fatal, are sometimes referred to as spongiform encephalopathies. They are so named because they frequently cause the brain to become riddled with holes. These ills, which can brew for years (or even for decades in humans) are widespread in animals.

The most common form is scrapie, found in sheep and goats. Afflicted animals lose coordination and eventually become so incapacitated that they cannot stand. They also become irritable and, in some cases, develop an intense itch that leads them to scrape off their wool or hair (hence the name "scrapie"). The other prion diseases of animals go by such names as transmissible mink encephalopathy, chronic wasting disease of mule deer and elk, feline spongiform encephalopathy and bovine spongiform encephalopathy. The last, often called mad cow disease, is the most worrisome.

Gerald A. H. Wells and John W. Wilesmith of the Central Veterinary Laboratory in Weybridge, England, identified the condition in 1986, after it began striking cows in Great Britain, causing them to became uncoordinated and unusually apprehensive. The source of the emerging epidemic was soon traced to a food supplement that included meat and bone meal from dead sheep. The



DIMITRY SCHIDLOVSKY

SOURCE: John W. Wilesmith

CATTLE WERE INCINERATED to prevent them from spreading "mad cow disease." This disorder, which has afflicted more than 130,000 cattle in Great Britain since the mid-1980s (*graph*), is one of several fatal neurodegenerative diseases of animals and humans thought to be caused by prions—infectious proteins. Studies are assessing whether prion disease can be transmitted from cows to people through the ingestion of beef.

DAVID JACKSON

methods for processing sheep carcasses had been changed in the late 1970s. Where once they would have eliminated the scrapie agent in the supplement, now they apparently did not. The British government banned the use of animal-derived feed supplements in 1988, and the epidemic has probably peaked. Nevertheless, many people continue to worry that they will eventually fall ill as a result of having consumed tainted meat.

The human prion diseases are more obscure. Kuru has been seen only among the Fore highlanders of Papua New Guinea. They call it the "laughing death." Vincent Zigas of the Australian Public Health Service and D. Carleton Gajdusek of the U.S. National Institutes of Health described it in 1957, noting that many highlanders became afflicted with a strange, fatal disease marked by loss of coordination (ataxia) and often later by dementia. The affected individuals probably acquired kuru through ritual cannibalism: the Fore tribe reportedly honored the dead by eating their brains. The practice has since stopped, and kuru has virtually disappeared.

Creutzfeldt-Jakob disease, in contrast, occurs worldwide and usually becomes evident as dementia. Most of the time it appears sporadically, striking one person in a million, typically around age 60. About 10 to 15 percent of cases are inherited, and a small number are, sadly, iatrogenic—spread inadvertently by the attempt to treat some other medical problem. Iatrogenic Creutzfeldt-Jakob disease has apparently been transmitted by corneal transplantation, implantation of dura mater or electrodes in the brain, use of contaminated surgical instruments, and injection of growth hormone derived from human pituitaries (before recombinant growth hormone became available).

The two remaining human disorders are Gerstmann-Sträussler-Scheinker disease (which is manifest as ataxia and other signs of damage to the cerebellum) and fatal familial insomnia (in which dementia follows difficulty sleeping). Both these conditions are usually inherited and typically appear in midlife. Fatal familial insomnia was discovered only recently, by Elio Lugaresi and Rossella Medori of the University of Bologna and Pierluigi Gambetti of Case Western Reserve University.

STANLEY B. PRUSINER is professor of neurology and biochemistry at the University of California School of Medicine, San Francisco. He is a member of the National Academy of Sciences, the Institute of Medicine and the American Academy of Arts and Sciences. He has won many awards for his research into prions, most recently the Albert Lasker Basic Medical Research Award and the Paul Ehrlich Award. This is his second article for *Scientific American.*

In Search of the Cause

I first became intrigued by the prion diseases in 1972, when as a resident in neurology at the University of California School of Medicine at San Francisco, I lost a patient to Creutzfeldt-Jakob disease. As I reviewed the scientific literature on that and related conditions, I learned that scrapie, Creutzfeldt-Jakob

disease and kuru had all been shown to be transmissible by injecting extracts of diseased brains into the brains of healthy animals. The infections were thought to be caused by a slow-acting virus, yet no one had managed to isolate the culprit.

In the course of reading, I came across an astonishing report in which Tikvah Alper and her colleagues at the Hammersmith Hospital in London suggested that the scrapie agent might lack nucleic acid, which usually can be degraded by ultraviolet or ionizing radiation. When the nucleic acid in extracts of scrapie-infected brains was presumably destroyed by those treatments, the extracts retained their ability to transmit scrapie. If the organism did lack DNA and RNA, the finding would mean that it was not a virus or any other known type of infectious agent, all of which contain genetic material. What, then, was it? Investigators had many ideas—including, jokingly, linoleum and kryptonite—but no hard answers.

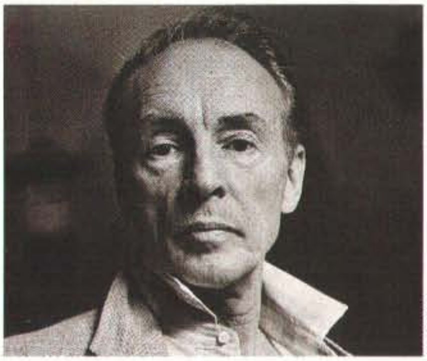
I immediately began trying to solve this mystery when I set up a laboratory at U.C.S.F. in 1974. The first step had to be a mechanical one—purifying the infectious material in scrapie-infected brains so that its composition could be analyzed. The task was daunting; many investigators had tried and failed in the past. But with the optimism of youth, I forged ahead [see "Prions," by Stanley B. Prusiner; SCIENTIFIC AMERICAN, October 1984]. By 1982 my colleagues and I had made good progress, producing extracts of hamster brains consisting almost exclusively of infectious material. We had, furthermore, subjected the extracts to a range of tests designed to reveal the composition of the disease-causing component.

Amazing Discovery

All our results pointed toward one startling conclusion: the infectious agent in scrapie (and presumably in the related diseases) did indeed lack nucleic acid and consisted mainly, if not exclusively, of protein. We deduced that DNA and RNA were absent because, like Alper, we saw that procedures known to damage nucleic acid did not reduce infectivity. And we knew protein was an essential component because procedures that denature (unfold) or degrade protein reduced infectivity. I thus introduced the term "prion" to distinguish this class of disease conveyer from viruses, bacteria, fungi and other known pathogens. Not long afterward, we determined that scrapie prions contained a single protein that we called PrP, for "prion protein."

Now the major question became, Where did the instructions specifying the sequence of amino acids in PrP reside? Were they carried by an undetected piece of DNA that traveled with PrP, or were they, perhaps, contained in a gene housed in the chromosomes of cells? The key to this riddle was the identification in 1984 of some 15 amino acids at one end of the PrP protein. My group identified this short amino acid sequence in collaboration with Leroy E. Hood and his co-workers at the California Institute of Technology.

Knowledge of the sequence allowed us and others to construct molecular probes, or detectors, able to indicate whether mammalian cells carried the PrP gene. With probes produced by Hood's team, Bruno Oesch, working in the laboratory of Charles Weissmann at the University of Zurich, showed that hamster cells do contain a gene for PrP. At about the same time, Bruce Chese-



MAGNUM

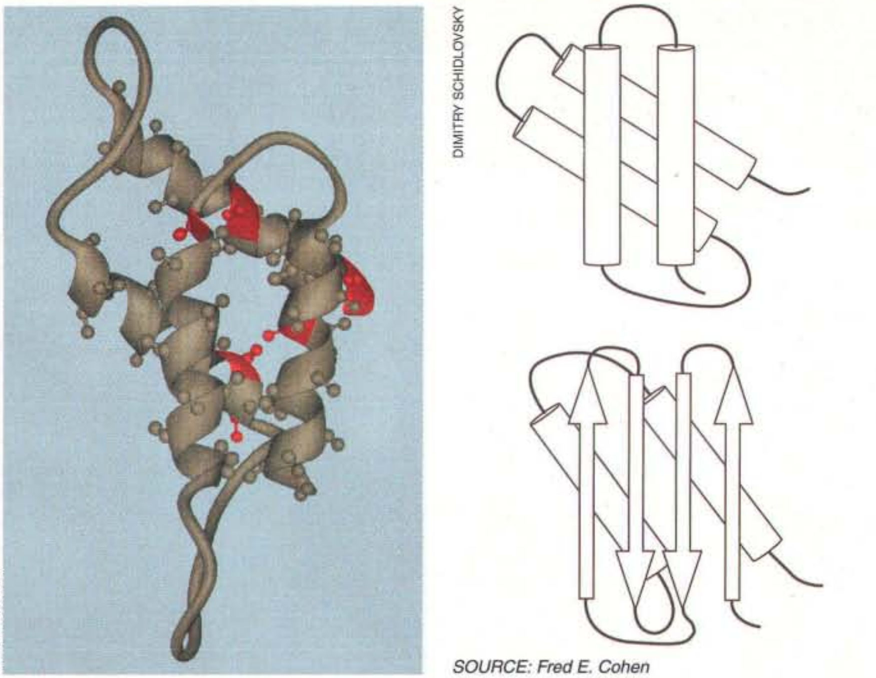
PRION DISEASES OF HUMANS (*table*), which may incubate for 30 years or more, can all cause progressive decline in cognition and motor function; hence, the distinctions among them are sometimes blurry. As the genetic mutations underlying familial forms of the diseases are found, those disorders are likely to be identified by their associated mutations alone. Choreographer George Balanchine (*photograph*) died of sporadic Creutzfeldt-Jakob disease in 1983 at age 79.

DISEASE	TYPICAL SYMPTOMS	ROUTE OF ACQUISITION	DISTRIBUTION	SPAN OF OVERT ILLNESS
Kuru	Loss of coordination, often followed by dementia	Infection (probably through cannibalism, which stopped by 1958)	Known only in highlands of Papua New Guinea; some 2,600 cases have been identified since 1957	Three months to one year
Creutzfeldt-Jakob disease	Dementia, followed by loss of coordination, although sometimes the sequence is reversed	Usually unknown (in "sporadic" disease)	*Sporadic form:* 1 person per million worldwide	Typically about one year; range is one month to more than 10 years
		Sometimes (in 10 to 15 percent of cases) inheritance of a mutation in the gene coding for the prion protein (PrP)	*Inherited form:* some 100 extended families have been identified	
		Rarely, infection (as an inadvertent consequence of a medical procedure)	*Infectious form:* about 80 cases have been identified	
Gerstmann-Sträussler-Scheinker disease	Loss of coordination, often followed by dementia	Inheritance of a mutation in the PrP gene	Some 50 extended families have been identified	Typically two to six years
Fatal familial insomnia	Trouble sleeping and disturbance of autonomic nervous system, followed by insomnia and dementia	Inheritance of a mutation in the PrP gene	Nine extended families have been identified	Typically about one year

boro of the NIH Rocky Mountain Laboratories made his own probes and established that mouse cells harbor the gene as well. That work made it possible to isolate the gene and to establish that it resides not in prions but in the chromosomes of hamsters, mice, humans and all other mammals that have been examined. What is more, most of the time, these animals make PrP without getting sick.

One interpretation of such findings was that we had made a terrible mistake: PrP had nothing to do with prion diseases. Another possibility was that PrP could be produced in two forms, one that generated disease and one that did not. We soon showed the latter interpretation to be correct.

The critical clue was the fact that the PrP found in infected brains resisted breakdown by cellular enzymes called proteases. Most proteins in cells are degraded fairly easily. I therefore suspected that if a normal, nonthreatening form of PrP existed, it too would be susceptible to degradation. Ronald A. Barry in my laboratory then identified this hypothetical protease-sensitive form. It thus became clear that scrapie-causing PrP is a variant of a normal protein. We therefore called the normal protein "cellular PrP" and the infectious (protease-resistant) form "scrapie PrP." The latter term is now used to refer to the protein molecules that constitute the prions causing all scrapielike diseases of animals and humans.



FRED E. COHEN

DIMITRY SCHIDLOVSKY

SOURCE: Fred E. Cohen

PRION PROTEIN (PrP) is usually harmless. In its benign state, its backbone twists into multiple helices (shown as spirals in the plausible ribbon model at the left and as cylinders in the cartoon at the top right). PrP becomes the infectious, "scrapie" form—a prion—when much of the backbone stretches out, forming so-called beta strands (arrows in the hypothetical structure at the bottom right). Red sites on the ribbon model of normal PrP highlight positions at which substitution of one amino acid for another probably promotes folding into the scrapie form.

Prion Diseases Can Be Inherited

Early on we had hoped to use the PrP gene to generate pure copies of PrP. Next, we would inject the protein molecules into animals, secure in the knowledge that no elusive virus was clinging to them. If the injections caused scrapie in the animals, we would have shown that protein molecules could, as we had proposed, transmit disease. By 1986, however, we knew the plan would not work. For one thing, it proved very difficult to induce the gene to make the high levels of PrP needed for conducting studies. For another thing, the protein that was produced was the normal, cellular form. Fortunately, work on a different problem led us to an alternative approach for demonstrating that prions could transmit scrapie without the help of any accompanying nucleic acid.

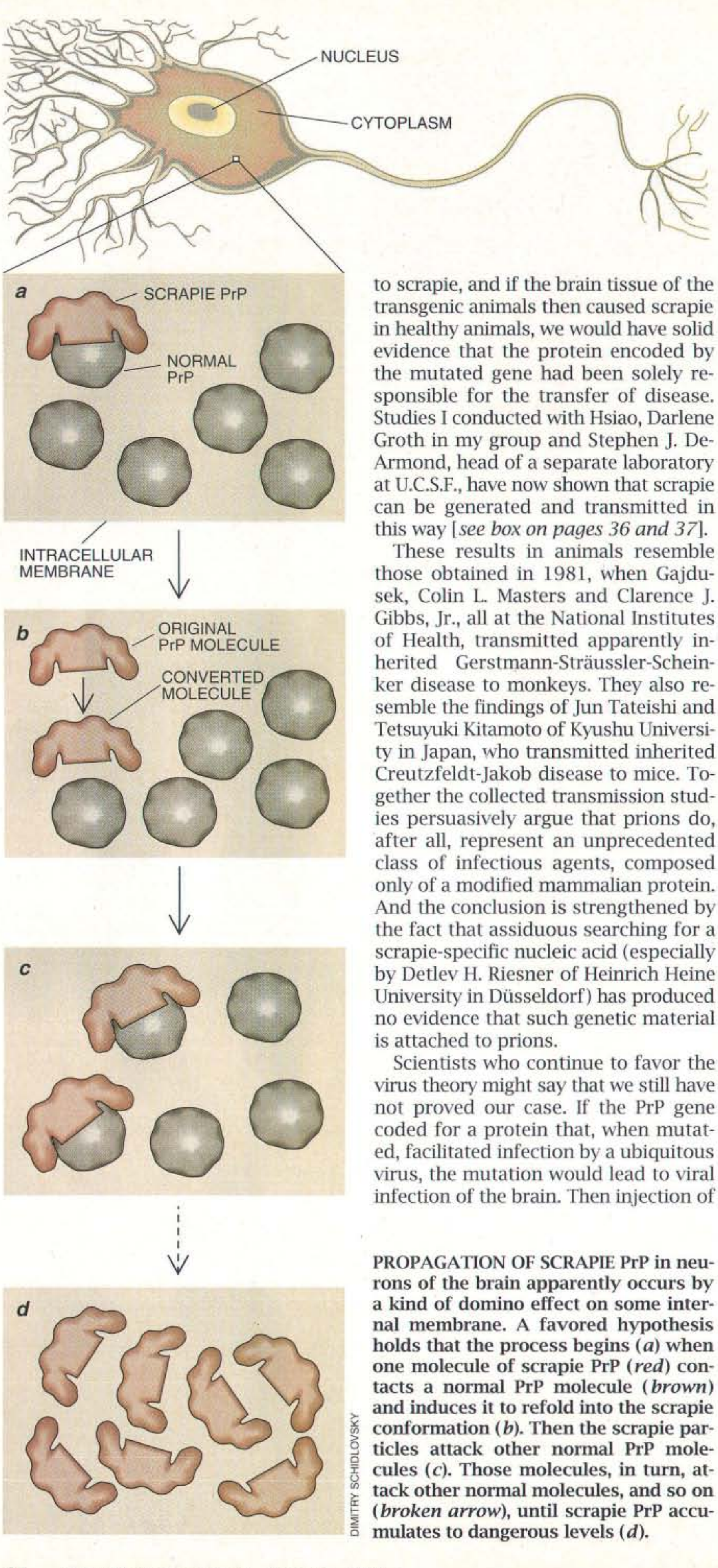
In many cases, the scrapielike illnesses of humans seemed to occur without having been spread from one host to another, and in some families they appeared to be inherited. (Today researchers know that about 10 percent of human prion diseases are familial, felling half of the members of the affected families.) It was this last pattern that drew our attention. Could it be that prions were more unusual than we originally thought? Were they responsible for the appearance of both hereditary and transmissible illnesses?

In 1988 Karen Hsiao in my laboratory and I uncovered some of the earliest data showing that human prion diseases can certainly be inherited. We acquired clones of a PrP gene obtained from a man who had Gerstmann-Sträussler-Scheinker disease in his family and was dying of it himself. Then we compared his gene with PrP genes obtained from a healthy population and found a tiny abnormality known as a point mutation.

To grasp the nature of this mutation, it helps to know something about the organization of genes. Genes consist of two strands of the DNA building blocks called nucleotides, which differ from one another in the bases they carry. The bases on one strand combine with the bases on the other strand to form base pairs: the "rungs" on the familiar DNA "ladder." In addition to holding the DNA ladder together, these pairs spell out the sequence of amino acids that must be strung together to make a particular protein. Three base pairs together—a unit called a codon—specify a single amino acid. In our dying patient, just one base pair (out of more than 750) had been exchanged for a different pair. The change, in turn, had altered the information carried by codon 102, causing the amino acid leucine to be substituted for the amino acid proline in the man's PrP protein.

With the help of Tim J. Crow of Northwick Park Hospital in London and Jurg Ott of Columbia University and their colleagues, we discovered the same mutation in genes from a large number of patients with Gerstmann-Sträussler-Scheinker disease, and we showed that the high incidence in the affected families was statistically significant. In other words, we established genetic linkage between the mutation and the disease—a finding that strongly implies the mutation is the cause. Over the past six years work by many investigators has uncovered 18 mutations in families with inherited prion diseases; for five of these mutations, enough cases have now been collected to demonstrate genetic linkage.

The discovery of mutations gave us a way to eliminate the possibility that a nucleic acid was traveling with prion proteins and directing their multiplication. We could now create genetically altered mice carrying a mutated PrP gene. If the presence of the altered gene in these "transgenic" animals led by itself



PROPAGATION OF SCRAPIE PrP in neurons of the brain apparently occurs by a kind of domino effect on some internal membrane. A favored hypothesis holds that the process begins (*a*) when one molecule of scrapie PrP (*red*) contacts a normal PrP molecule (*brown*) and induces it to refold into the scrapie conformation (*b*). Then the scrapie particles attack other normal PrP molecules (*c*). Those molecules, in turn, attack other normal molecules, and so on (*broken arrow*), until scrapie PrP accumulates to dangerous levels (*d*).

DIMITRY SCHIDLOVSKY

to scrapie, and if the brain tissue of the transgenic animals then caused scrapie in healthy animals, we would have solid evidence that the protein encoded by the mutated gene had been solely responsible for the transfer of disease. Studies I conducted with Hsiao, Darlene Groth in my group and Stephen J. DeArmond, head of a separate laboratory at U.C.S.F., have now shown that scrapie can be generated and transmitted in this way [*see box on pages 36 and 37*].

These results in animals resemble those obtained in 1981, when Gajdusek, Colin L. Masters and Clarence J. Gibbs, Jr., all at the National Institutes of Health, transmitted apparently inherited Gerstmann-Sträussler-Scheinker disease to monkeys. They also resemble the findings of Jun Tateishi and Tetsuyuki Kitamoto of Kyushu University in Japan, who transmitted inherited Creutzfeldt-Jakob disease to mice. Together the collected transmission studies persuasively argue that prions do, after all, represent an unprecedented class of infectious agents, composed only of a modified mammalian protein. And the conclusion is strengthened by the fact that assiduous searching for a scrapie-specific nucleic acid (especially by Detlev H. Riesner of Heinrich Heine University in Düsseldorf) has produced no evidence that such genetic material is attached to prions.

Scientists who continue to favor the virus theory might say that we still have not proved our case. If the PrP gene coded for a protein that, when mutated, facilitated infection by a ubiquitous virus, the mutation would lead to viral infection of the brain. Then injection of brain extracts from the mutant animal would spread the infection to another host. Yet in the absence of any evidence of a virus, this hypothesis looks to be untenable.

In addition to showing that a protein can multiply and cause disease without help from nucleic acids, we have gained insight into how scrapie PrP propagates in cells. Many details remain to be worked out, but one aspect appears quite clear: the main difference between normal PrP and scrapie PrP is conformational. Evidently, the scrapie protein propagates itself by contacting normal PrP molecules and somehow causing them to unfold and flip from their usual conformation to the scrapie shape. This change initiates a cascade in which newly converted molecules change the shape of other normal PrP molecules, and so on. These events apparently occur on a membrane in the cell interior.

We started to think that the differences between cellular and scrapie forms of PrP must be conformational after other possibilities began to seem unlikely. For instance, it has long been known that the infectious form often has the same amino acid sequence as the normal type. Of course, molecules that start off being identical can later be chemically modified in ways that alter their activity. But intensive investigations by Neil Stahl and Michael A. Baldwin in my laboratory have turned up no differences of this kind.

One Protein, Two Shapes

How, exactly, do the structures of normal and scrapie forms of PrP differ? Studies by Keh-Ming Pan in our group indicate that the normal protein consists primarily of alpha helices, regions in which the protein backbone twists into a specific kind of spiral; the scrapie form, however, contains beta strands, regions in which the backbone is fully extended. Collections of these strands form beta sheets. Fred E. Cohen, who directs another laboratory at U.C.S.F., has used molecular modeling to try to predict the structure of the normal protein based on its amino acid sequence. His calculations imply that the protein probably folds into a compact structure having four helices in its core. Less is known about the structure, or structures, adopted by scrapie PrP.

The evidence supporting the proposition that scrapie PrP can induce an alpha-helical PrP molecule to switch to a beta-sheet form comes primarily from two important studies by investigators in my group. María Gasset learned that synthetic peptides (short strings of amino acids) corresponding to three of the four putative alpha-helical regions of PrP can fold into beta sheets. And Jack Nguyen has shown that in their beta-sheet conformation, such peptides can impose a beta-sheet structure on helical PrP peptides. More recently Byron W. Caughey of the Rocky Mountain Laboratories and Peter T. Lansbury of the Massachusetts Institute of Technology have reported that cellular PrP can be converted into scrapie PrP in a test tube by mixing the two proteins together.

PrP molecules arising from mutated genes probably do not adopt the scrapie conformation as soon as they are synthesized. Otherwise, people carrying mutant genes would become sick in early childhood. We suspect that mutations in the PrP gene render the resulting proteins susceptible to flipping from an alpha-helical to a beta-sheet shape. Presumably, it takes time until one of the molecules spontaneously flips over and still more time for scrapie PrP to accumulate and damage the brain enough to cause symptoms.

Fred Cohen and I think we might be able to explain why the various mutations that have been noted in PrP genes could facilitate folding into the beta-sheet form. Many of the human mutations give rise to the substitution of one amino acid for another within the four putative helices or at their borders. Insertion of incorrect amino acids at those positions might destabilize a helix, thus increasing the likelihood that the affected helix and its neighbors will refold into a beta-sheet conformation. Conversely, Hermann Schätzel in my laboratory finds that the harmless differences distinguishing the PrP gene of humans from those of apes and monkeys affect amino acids lying outside of the proposed helical domains—where the divergent amino acids probably would not profoundly influence the stability of the helical regions.

Treatment Ideas Emerge

No one knows exactly how propagation of scrapie PrP damages cells. In cell cultures, the conversion of normal PrP to the scrapie form occurs inside neurons, after which scrapie PrP accumulates in intracellular vesicles known as lysosomes. In the brain, filled lysosomes could conceivably burst and damage cells. As the diseased cells died, creating holes in the brain, their prions would be released to attack other cells.

We do know with certainty that cleavage of scrapie PrP is what produces PrP fragments that accumulate as plaques in the brains of some patients. Those aggregates resemble plaques seen in Alzheimer's disease, although the Alzheimer's clumps consist of a different protein. The PrP plaques are a useful sign of prion infection, but they seem not to be a major cause of impairment. In many people and animals with prion disease, the plaques do not arise at all.

Even though we do not yet know much about how PrP scrapie harms brain tissue, we can foresee that an understanding of the three-dimensional structure of the PrP protein will lead to therapies. If, for example, the four-helix-bundle model of PrP is correct, drug developers might be able to design a compound that would bind to a central pocket that could be formed by the four helices. So bound, the drug would stabilize these helices and prevent their conversion into beta sheets.

Another idea for therapy is inspired by research in which Weissmann and his colleagues applied gene-targeting technology to create mice that lacked the PrP gene and so could not make PrP. By knocking out a gene and noting the consequences of its loss, one can often deduce the usual functions of the gene's protein product. In this case, however, the animals missing PrP displayed no detectable abnormalities. If it turns out that PrP is truly inessential, then physicians might one day consider delivering so-called antisense or antigene therapies to the brains of patients with prion diseases. Such therapies aim to block genes from giving rise to unwanted proteins and could potentially shut down production of cellular PrP [see "The New Genetic Medicines," by Jack S. Cohen and Michael E. Hogan; SCIENTIFIC AMERICAN, December 1994]. They would thereby block PrP from propagating itself.

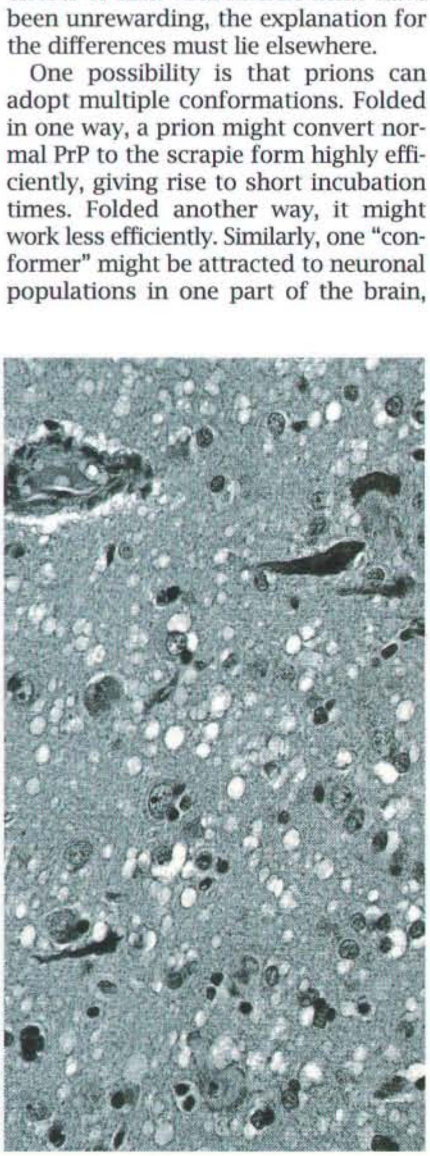
It is worth noting that the knockout mice provided a welcomed opportunity to challenge the prion hypothesis. If the animals became ill after inoculation with prions, their sickness would have indicated that prions could multiply even in the absence of a preexisting pool of PrP molecules. As I expected, inoculation with prions did not produce scrapie, and no evidence of prion replication could be detected.

The enigma of how scrapie PrP multiplies and causes disease is not the only puzzle starting to be solved. Another long-standing question—the mystery of how prions consisting of a single kind of protein can vary markedly in their effects—is beginning to be answered as well. Iain H. Pattison of the Agriculture Research Council in Compton, England, initially called attention to this phenomenon. Years ago he obtained prions from two separate sets of goats. One isolate made inoculated animals drowsy, whereas the second made them hyperactive. Similarly, it is now evident that some prions cause disease quickly, whereas others do so slowly.

The Mystery of "Strains"

Alan G. Dickinson, Hugh Fraser and Moira E. Bruce of the Institute for Animal Health in Edinburgh, who have examined the differential effects of varied isolates in mice, are among those who note that only pathogens containing nucleic acids are known to occur in multiple strains. Hence, they and others assert, the existence of prion "strains" indicates the prion hypothesis must be incorrect; viruses must be at the root of scrapie and its relatives. Yet because efforts to find viral nucleic acids have been unrewarding, the explanation for the differences must lie elsewhere.

One possibility is that prions can adopt multiple conformations. Folded in one way, a prion might convert normal PrP to the scrapie form highly efficiently, giving rise to short incubation times. Folded another way, it might work less efficiently. Similarly, one "conformer" might be attracted to neuronal populations in one part of the brain,



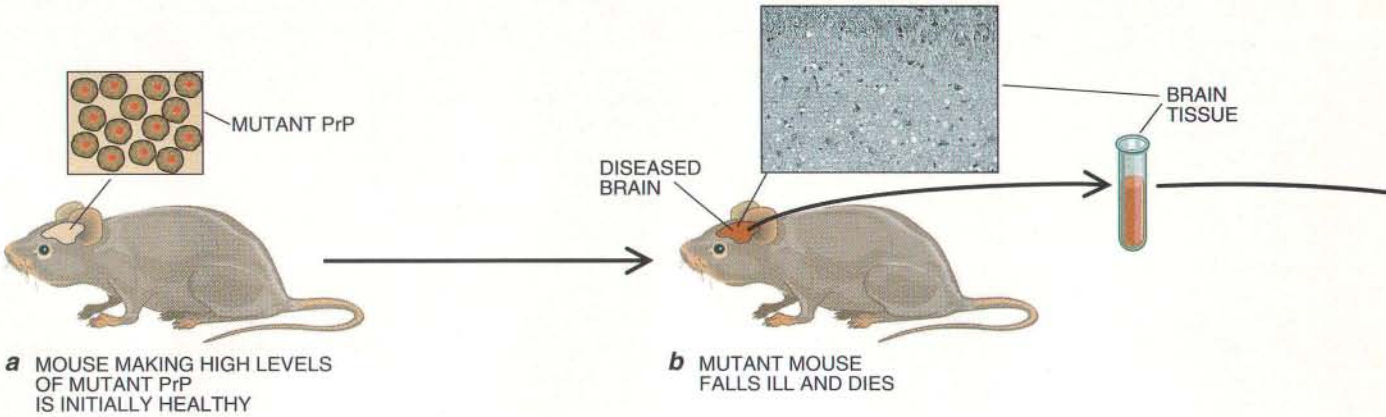
HOLES IN BRAIN TISSUE (*white spots*) are a frequent feature of prion diseases. They give the brain a spongelike appearance. This micrograph shows the cerebral cortex of a patient suffering from Creutzfeldt-Jakob disease.

STEPHEN J. DEARMOND

A Persuasive Experiment

Several studies have shown that prions composed only of PrP are able to convey infection from one animal to another. In one such experiment, the author and his colleagues created mice carrying many copies of a mutant PrP gene (*a*); these animals made high levels of mutant PrP, some of which appears to adopt the scrapie conformation. Eventually all the mice displayed symptoms of brain damage and died (*b*). Then the workers injected brain tissue from the diseased animals into genetically altered mice making low levels of the same mutant PrP protein. (Such

DIMITRY SCHIDLOVSKY (drawings); STEPHEN J. DeARMOND (micrographs)


MUTANT PrP
DISEASED BRAIN
BRAIN TISSUE
a MOUSE MAKING HIGH LEVELS OF MUTANT PrP IS INITIALLY HEALTHY
b MUTANT MOUSE FALLS ILL AND DIES

whereas another might be attracted to neurons elsewhere, thus producing different symptoms. Considering that PrP can fold in at least two ways, it would not be surprising to find it can collapse into other structures as well.

Since the mid-1980s we have also sought insight into a phenomenon known as the species barrier. This concept refers to the fact that something makes it difficult for prions made by one species to cause disease in animals of another species. The cause of this difficulty is of considerable interest today because of the epidemic of mad cow disease in Britain. We and others have been trying to find out whether the species barrier is strong enough to prevent the spread of prion disease from cows to humans.

Breaking the Barrier

The barrier was discovered by Pattison, who in the 1960s found it hard to transmit scrapie between sheep and rodents. To determine the cause of the trouble, my colleague Michael R. Scott and I later generated transgenic mice expressing the PrP gene of the Syrian hamster—that is, making the hamster PrP protein. The mouse gene differs from that of the hamster gene at 16 codons out of 254. Normal mice inoculated with hamster prions rarely acquire scrapie, but the transgenic mice became ill within about two months.

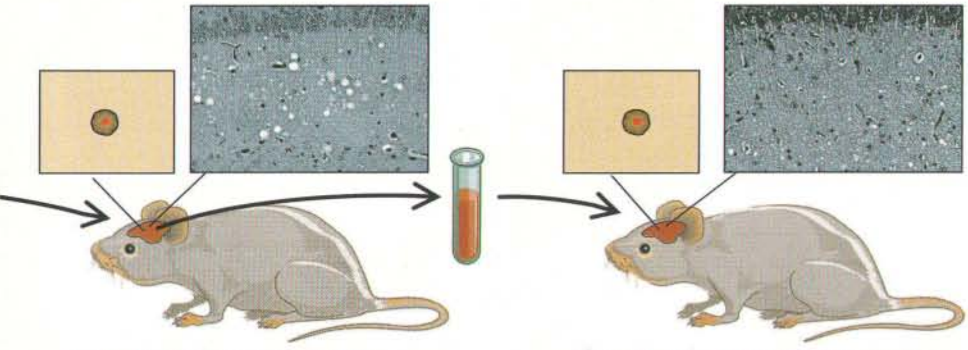
We thus concluded that we had broken the species barrier by inserting the hamster genes into the mice. Moreover, on the basis of this and other experiments, we realized that the barrier resides in the amino acid sequence of PrP: the more the sequence of a scrapie PrP molecule resembles the PrP sequence of its host, the more likely it is that the host will acquire prion disease. In one of those other experiments, for example, we examined transgenic mice carrying the Syrian hamster PrP gene in addition to their own mouse gene. Those mice make normal forms of both hamster and mouse PrP. When we inoculated the animals with mouse prions, they made more mouse prions. When we inoculated them with hamster prions, they made hamster prions. From this behavior, we learned that prions preferentially interact with cellular PrP of homologous, or like, composition.

The attraction of scrapie PrP for cellular PrP having the same sequence probably explains why scrapie managed to spread to cows in England from food consisting of sheep tissue: sheep and bovine PrP differ only at seven positions. In contrast, the sequence difference between human and bovine PrP is large: the molecules diverge at more than 30 positions. Because the variance is great, the likelihood of transmission from cows to people would seem to be low. Consistent with this assessment are epidemiological studies by W. Bryan Matthews, a professor emeritus at the University of Oxford. Matthews found no link between scrapie in sheep and the occurrence of Creutzfeldt-Jakob disease in sheep-farming countries.

On the other hand, two farmers who had "mad cows" in their herds have recently died of Creutzfeldt-Jakob disease. Their deaths may have nothing to do with the bovine epidemic, but the situation bears watching. It may turn out that certain parts of the PrP molecule are more important than others for breaking the species barrier. If that is the case, and if cow PrP closely resembles human PrP in the critical regions, then the likelihood of danger might turn out to be higher than a simple comparison of the complete amino acid sequences would suggest.

We began to consider the possibility that some parts of the PrP molecule might be particularly important to the species barrier after a study related to this blockade took an odd turn. My colleague Glenn C. Telling had created transgenic mice carrying a hybrid PrP gene that consisted of human codes flanked on either side by mouse codes; this gene gave rise to a hybrid protein. Then he introduced brain tissue from patients who had died of Creutzfeldt-Jakob disease or Gerstmann-Sträussler-Scheinker disease into the transgenic animals. Oddly enough, the animals became ill much more frequently and faster than did mice carrying a full human PrP gene, which diverges from mouse PrP at 28 positions. This outcome implied that similarity in the central region of the PrP molecule may be more critical than it is in the other segments.

The result also lent support to earlier indications—uncovered by Shu-Lian Yang in DeArmond's laboratory and Albert Taraboulos in my group—that molecules made by the host can influence the behavior of scrapie PrP. We speculate that in the hybrid-gene study, a mice were chosen as recipients because scrapie PrP is most attracted to PrP molecules having the same composition.) Uninoculated mice did not become ill (indicating that making low levels of the aberrant protein was safe), but many of the treated ones did (*c*). Moreover, brain tissue transferred from the diseased recipients to their healthy counterparts caused illness once again (*d*). If the aberrant protein were unable to transmit infection, none of the inoculated animals would have sickened.



c MOUSE MAKING LOW LEVELS OF MUTANT PrP BECOMES ILL AFTER INOCULATION

d IDENTICAL MOUSE BECOMES ILL AFTER RECEIVING INOCULATION FROM FIRST RECIPIENT

mouse protein, possibly a "chaperone" normally involved in folding nascent protein chains, recognized one of the two mouse-derived regions of the hybrid PrP protein. This chaperone bound to that region and helped to refold the hybrid molecule into the scrapie conformation. The chaperone did not provide similar help in mice making a totally human PrP protein, presumably because the human protein lacked a binding site for the mouse factor.

The List May Grow

An unforeseen story has recently emerged from studies of transgenic mice making unusually high amounts of normal PrP proteins. DeArmond, David Westaway in our group and George A. Carlson of the McLaughlin Laboratory in Great Falls, Mont., became perplexed when they noted that some older transgenic mice developed an illness characterized by rigidity and diminished grooming. When we pursued the cause, we found that making excessive amounts of PrP can eventually lead to neurodegeneration and, surprisingly, to destruction of both muscles and peripheral nerves. These discoveries widen the spectrum of prion diseases and are prompting a search for human prion diseases that affect the peripheral nervous system and muscles.

Investigations of animals that overproduce PrP have yielded another benefit as well. They offer a clue as to how the sporadic form of Creutzfeldt-Jakob disease might arise. For a time I suspected that sporadic disease might begin when the wear and tear of living led to a mutation of the PrP gene in at least one cell in the body. Eventually, the mutated protein might switch to the scrapie form and gradually propagate itself, until the buildup of scrapie PrP crossed the threshold to overt disease. The mouse studies suggest that at some point in the lives of the one in a million individuals who acquire sporadic Creutzfeldt-Jakob disease, cellular PrP may spontaneously convert to the scrapie form. The experiments also raise the possibility that people who become afflicted with sporadic Creutzfeldt-Jakob disease overproduce PrP, but we do not yet know if, in fact, they do.

All the known prion diseases in humans have now been modeled in mice. With our most recent work we have inadvertently developed an animal model for sporadic prion disease. Mice inoculated with brain extracts from scrapie-infected animals and from humans afflicted with Creutzfeldt-Jakob disease have long provided a model for the infectious forms of prion disorders. And the inherited prion diseases have been modeled in transgenic mice carrying mutant PrP genes. These murine representations of the human prion afflictions should not only extend understanding of how prions cause brain degeneration, they should also create opportunities to evaluate therapies for these devastating maladies.

Striking Similarities

Ongoing research may also help determine whether prions consisting of other proteins play a part in more common neurodegenerative conditions, including Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis. There are some marked similarities in all these disorders. As is true of the known prion diseases, the more widespread ills mostly occur sporadically but sometimes "run" in families. All are also usually diseases of middle to later life and are marked by similar pathology: neurons degenerate, protein deposits can accumulate as plaques, and glial cells (which support and nourish nerve cells) grow larger in reaction to damage to neurons. Strikingly, in none of these disorders do white blood cells—those ever present warriors of the immune system—infiltrate the brain. If a virus were involved in these illnesses, white cells would be expected to appear.

Recent findings in yeast encourage speculation that prions unrelated in amino acid sequence to the PrP protein could exist. Reed B. Wickner of the NIH reports that a protein called Ure2p might sometimes change its conformation, thereby affecting its activity in the cell. In one shape, the protein is active; in the other, it is silent.

The collected studies described here argue persuasively that the prion is an entirely new class of infectious pathogen and that prion diseases result from aberrations of protein conformation. Whether changes in protein shape are responsible for common neurodegenerative diseases, such as Alzheimer's, remains unknown, but it is a possibility that should not be ignored.

FURTHER READING

SCRAPIE DISEASE IN SHEEP. Herbert B. Parry. Edited by D. R. Oppenheimer. Academic Press, 1983.

MOLECULAR BIOLOGY OF PRION DISEASES. S. B. Prusiner in *Science*, Vol. 252, pages 1515–1522; June 14, 1991.

PRION DISEASES OF HUMANS AND ANIMALS. Edited by S. B. Prusiner, J. Collinge, J. Powell and B. Anderton. Ellis Horwood, 1992.

FATAL FAMILIAL INSOMNIA: INHERITED PRION DISEASES, SLEEP, AND THE THALAMUS. Edited by C. Guilleminault et al. Raven Press, 1994.

MOLECULAR BIOLOGY OF PRION DISEASES. Special issue of *Philosophical Transactions of the Royal Society of London, Series B*, Vol. 343, No. 1306; March 29, 1994.

STRUCTURAL CLUES TO PRION REPLICATION. F. E. Cohen, K.-M. Pan, Z. Huang, M. Baldwin, R. J. Fletterick and S. B. Prusiner in *Science*, Vol. 264, pages 530–531; April 22, 1994.

Earth before Pangea

The North American continent may be more nomadic than any of its inhabitants

by Ian W. D. Dalziel

The little airplane banked to the right. From my seat on the port side I could see its shadow crossing the ice. The skis made it look rather like a duck coming in to land on water, webbed feet outstretched. As the pilot leveled the aircraft, a huge cliff came into view, the dark brown of its rocks contrasting sharply with the pristine whiteness of ice and snow that faded into the horizon.

The steeply inclined layers of this Precambrian sandstone were distorted by concertinalike folds. I took several photographs. As we rounded the cliff, another came into view. Resting on top of the sandstone was a thin capping of rock almost as white as the background: Cambrian limestone. "Fascinating," I thought as I raised my camera again. "The basic geology here is very similar to that of western North America."

My colleagues and I had come to the Pensacola Mountains of Antarctica to study how the two geologic subdivisions—East and West—of the icy continent relate to each other. East Antarctica is an old Precambrian shield lying to the south of Australia, India and Africa; West Antarctica is part of the geologically young and active volcanic "ring of fire" that surrounds the Pacific Ocean. The uplifted rim of the East Antarctic shield meets West Antarctica along the Transantarctic Mountains, of which the Pensacolas form a northern extension.

It had been a long trip down: 14 hours from Los Angeles to New Zealand in a commercial jet, 10 hours from New Zealand to McMurdo Station in Antarctica in a ski-equipped Hercules transport and, finally, five hours across the continent to the Pensacola Mountains, bypassing the South Pole en route. Now, after setting up our base camp, we were finally at the mountains near the southern margin of the same ocean that laps the beaches of Los Angeles.

IAN W. D. DALZIEL has been studying the geology of Antarctica, the Andes, the Caledonides and the Canadian Shield since earning his Ph.D. at the University of Edinburgh in 1963. Currently he is a senior research scientist at the Institute for Geophysics of the University of Texas at Austin. In 1992 Dalziel received the Geological Society of London's Murchison Medal. When possible, he likes to travel with his wife, Linda, and their son and daughter, Kyle and Kacie, and to scull on Town Lake.

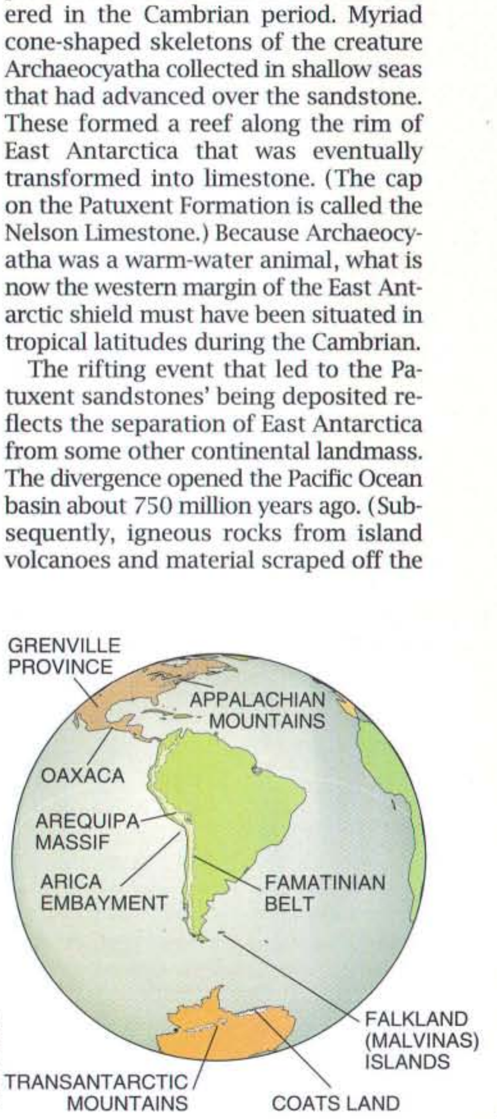
We still had to get to the rocks, however. In Antarctica such excursions take time. Having selected a possible crevasse-free landing site, our pilot brought the Twin Otter down for a "ski drag." That is, he put some weight on the landing gear but maintained enough airspeed to take off again. We circled and carefully examined these tracks. Crevasses can be hidden under snow, but here there were no telltale signs of blue cracks. Coming around again, we touched down and stopped quickly so as to reduce the chance of hitting rough ice beneath the snow. It was a bumpy landing, nonetheless, although the aircraft appeared to have suffered only superficial damage. We roped ourselves together for safety and started to walk across the windblown snow to the base of the cliff, leaving our anxious pilot to examine the plane.

Fossil Clues

The boundary between the two rock types exposed in the Pensacola Mountains is one of the most fundamental in the earth's history. After the birth of the planet 4.5 billion years ago came the four-billion-year-long interval of time known as the Precambrian. Toward the end of this era—about 750 million years ago, while the first soft-bodied, multicellular creatures were developing—the brown sandstones of the underlying Patuxent Formation we had just sighted were deposited. The strata were laid down in a rift valley that opened within the continental shield. As the rift deepened, rivers poured in, dropping their eroded soils onto the valley floor.

About 540 million years ago, an explosion of multicellular animal life ushered in the Cambrian period. Myriad cone-shaped skeletons of the creature Archaeocyatha collected in shallow seas that had advanced over the sandstone. These formed a reef along the rim of East Antarctica that was eventually transformed into limestone. (The cap on the Patuxent Formation is called the Nelson Limestone.) Because Archaeocyatha was a warm-water animal, what is now the western margin of the East Antarctic shield must have been situated in tropical latitudes during the Cambrian.

The rifting event that led to the Patuxent sandstones' being deposited reflects the separation of East Antarctica from some other continental landmass. The divergence opened the Pacific Ocean basin about 750 million years ago. (Subsequently, igneous rocks from island volcanoes and material scraped off the



IAN WORPOLE

GALEN ROWELL *Mountain Light*

TRANSANTARCTIC MOUNTAINS mark an ancient boundary between East Antarctica and another continent, probably North America. The Dry Valleys (*right*) are cut into the uplifted margin of the chain. Features on today's earth (*above*) record the travels of North America around other continents.

subducting ocean floor accreted onto East Antarctica, forming West Antarctica.) This rifting occurred long before the supercontinent Pangea—from which the present continents broke off—was formed. Pangea was assembled only at the end of the Paleozoic era, approximately 250 million years ago. It started to fragment during the Jurassic period of the Mesozoic era, approximately 170 million years ago, creating the Atlantic and other young ocean basins.

Making our way up a ridge toward the top of the cliff, we saw that the lowest layers of the Cambrian strata—which lie below the limestone—were made of pink conglomerate and coarse sandstones. As the sea advanced over the deepening rift and the subsiding margin, it had ground the Precambrian rocks into boulders, pebbles and sand grains. The deposits became more fine-grained as we climbed, and the quartz sandstones immediately underneath the Nelson Limestone had the appearance of old friends. They were full of vertical worm burrows known as *Skolithus.*

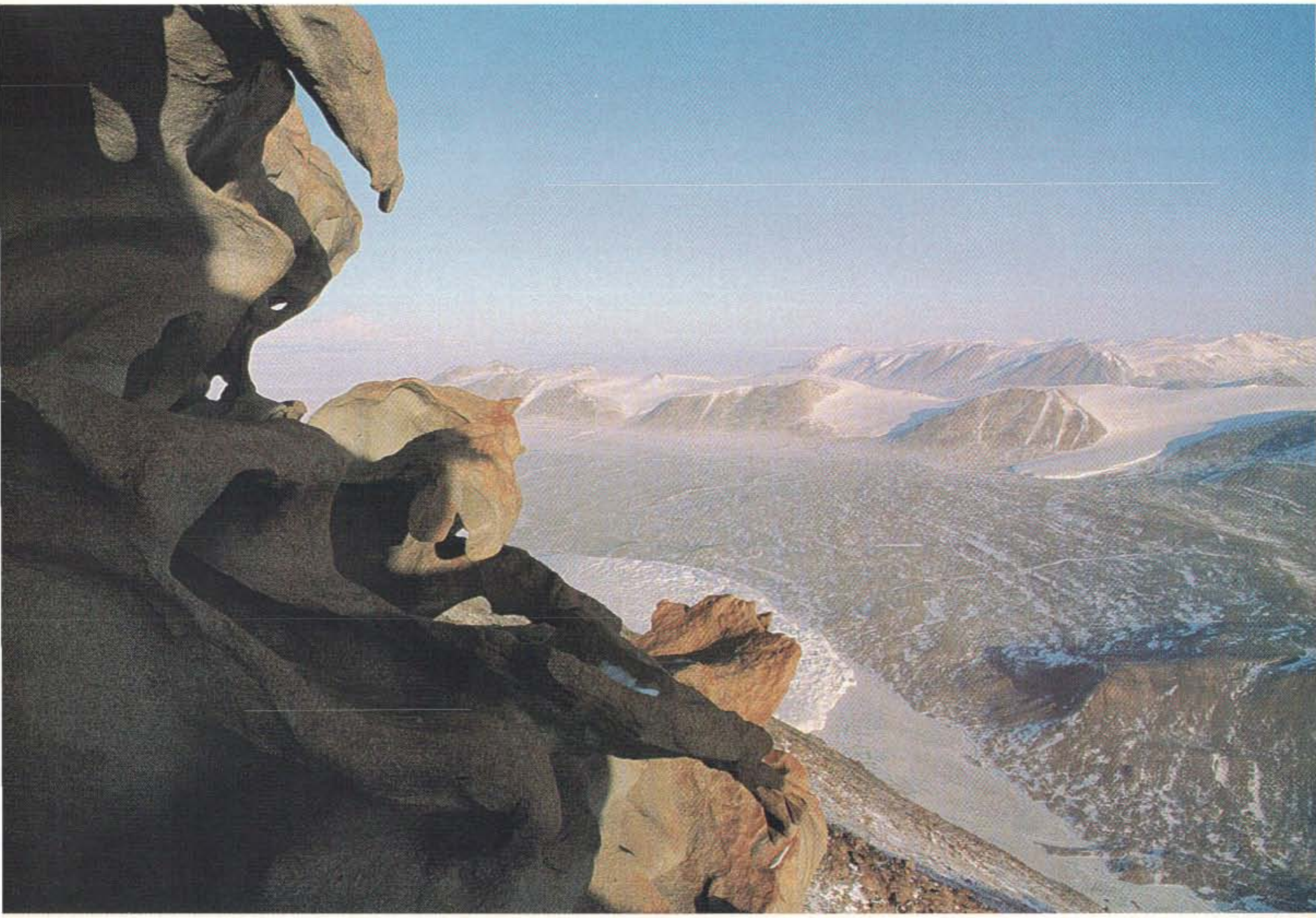
These tubes are the only traces of ancient filter-feeders, which extracted nutrients from sediments and left a clayey residue around their burrows. "Just like western North America," I noted out loud, "but then just like the Durness rocks of northwestern Scotland, too." Indeed, strata deposited by the seawater that advanced to cover most of the continents 540 million years ago—as evinced by the presence of Cambrian seashores in such places as Wisconsin—are remarkably similar on all continents.

Matching Mountains...

There is nothing like personal experience with rocks, however, to set a geologist thinking. My first impressions of the Transantarctic Mountains in 1987 raised a question that stayed near the forefront of my mind: Could the continent from which Antarctica rifted apart at the end of the Precambrian possibly have been western North America? Or were their margins at that distant time merely in similar environments on either side of an even more ancient Pacific Ocean basin?

The answer has far-reaching implications. The global paleogeography of the time ("paleo" is a prefix that geologists use to indicate "historical") is currently a mystery. To know how the continents were distributed could provide clues to the vast environmental alterations that preceded the Cambrian period. Late in Precambrian times there were several ice ages, and the oceanic and, presumably, atmospheric chemistry changed greatly. Multicellular animals evolved, heralding a biological profusion that included the far-distant ancestors of vertebrates and, hence, of human beings [see "End of the Proterozoic Eon," by Andrew H. Knoll; SCIENTIFIC AMERICAN, October 1991].

It is clearly difficult to map out, with any degree of certainty, the geography of an ancient time on a dynamic planet with continents that move. Alfred Wegener and other pioneers of the theory of continental drift had noted that several North and South American moun-



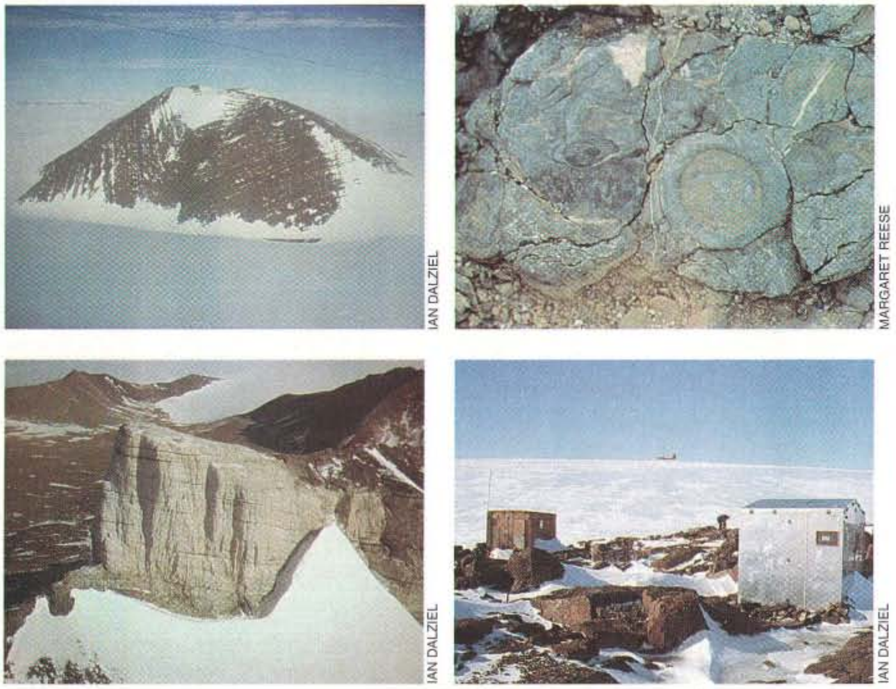
tain ranges truncated at the Atlantic margins match up neatly across the ocean with mountain ranges in Europe and Africa. Nowadays magnetic data and satellite images of the ocean floor showing fractures—appearing rather like railway tracks, along which the continents slid apart—allow us to reconstruct Pangea very accurately.

A number of lines of evidence indicate that Pangea was not the original configuration of the continents. When iron-bearing rocks solidify from lava, they become magnetized in the direction of the earth's magnetic field. The magnetization of rocks that congealed from pre-Mesozoic lava is quite different in North America and Africa, suggesting that in an earlier era these continents moved separately. Volcanic rocks that were fragments of ancient ocean floor have also been found in mountain ranges of Pangea such as the Famatinian belt (Argentina), the Mozambique belt (Africa) and the older Appalachians. These early Paleozoic and Precambrian ophiolites—as the rocks are called—demonstrate that former ocean basins closed when the supercontinent amalgamated. Struck in the 1960s by the presence of early Paleozoic ophiolites in the Appalachian Mountains of the Maritime Provinces in Canada, the imaginative Canadian geophysicist J. Tuzo Wilson asked: "Did the Atlantic Ocean open, close and then reopen?"

In reconstructing continental configurations prior to Pangea, we get no help from the ocean floors. Although the Pacific Ocean basin already existed, ocean floor of such antiquity has long been thrust under the continents bordering the basin. Geologists therefore have no oceanic "railway map" for continental drift before Pangea. We have to fall back on evidence from the continents themselves, just as Wegener did when trying to reconstruct Pangea before modern oceanography and satellites.

... and Margins

Within Pangea there are some ancient continental margins that have no obvious counterparts. The Pacific margins of North and South America, Antarctica and Australia were all formed near the end of the Precambrian, between 750 and 550 million years ago. The Appalachian margin of Laurentia—the ancestral shield of North America—also rifted away from another continent at that time. Since Wilson asked his famous question, the counterpart to this margin has usually been assumed to have been western Europe and northwestern Africa. But there is no firm evidence for such a juxtaposition.


IAN DALZIEL

MARGARET REESE

IAN DALZIEL

IAN DALZIEL

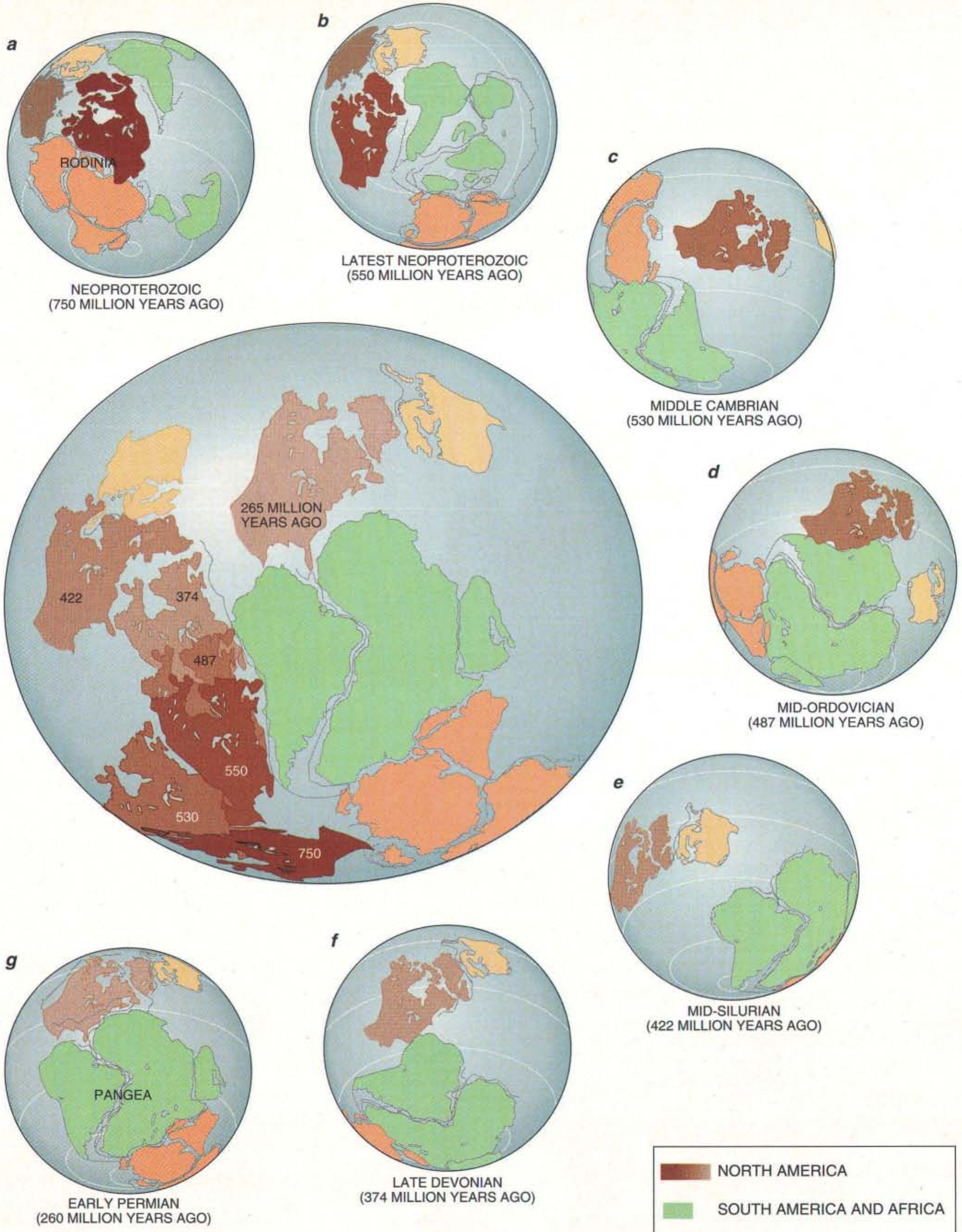
ROCK STRUCTURES in Antarctica provide clues to North America's voyages. Concertinalike folds (*top left*) in the Patuxent Formation mark the Precambrian boundary between North America and Antarctica. As the two continents rifted apart, Archaeocyatha (*top right*), among the first creatures with skeletons, formed a reef that fossilized into the Nelson Limestone cliff (*bottom left*). Outcroppings of rhyolite lavas in Littlewood Nunataks, Coats Land (*bottom right*), yield magnetic data that are being used to test the juxtaposition of North America and Antarctica. (The metal shack is a storm refuge for scientists from a nearby Argentine base.)

In 1989 I led another field trip to Antarctica, as part of the International Geological Congress hosted by the U.S. The object of the trip was to help bring Antarctic geology, long the private domain of a very small group of especially hardy souls (even among geologists!), into the mainstream of global earth science. Experts on the Himalayas, the European Alps, the Appalachians, the Rockies and many other regions participated.

Soon after, one of these scientists, Eldridge M. Moores, was browsing in the library of the University of California at Davis when he came across a short article by Richard T. Bell and Charles W. Jefferson of the Geological Survey of Canada. They pointed out similarities between Precambrian strata in western Canada and eastern Australia and concluded that the Pacific margins of Canada and Australia might have been juxtaposed. Sensitized by his recent trip, Moores realized this would imply that the Pacific margins of the U.S. and Antarctica had been juxtaposed, a thought similar to my own. After some quick library research, he sent me a map highlighting the structural parallels in the interiors of the Laurentian and East Antarctic shields. "Is this crazy?" he asked.

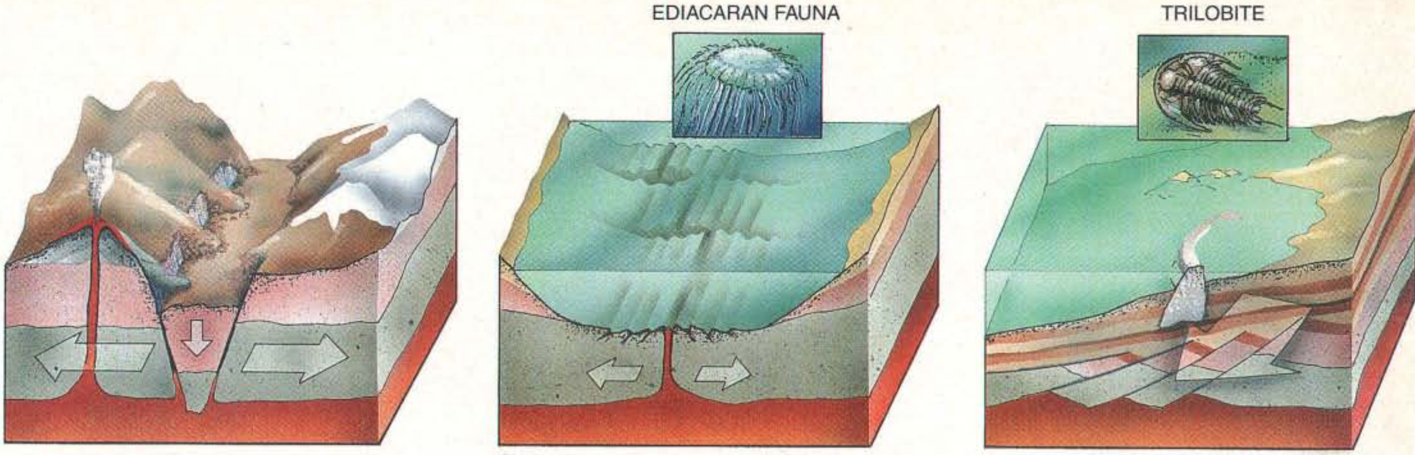
Similarities in the internal structures of displaced continents can be powerful evidence of former juxtaposition. Moores drew particular attention to a report citing that along the Transantarctic Mountains—in a place called the Shackleton Range (after the famous British explorer Sir Ernest Shackleton)—lie rocks similar in age and character to those underneath much of New Mexico and Arizona. He also pointed out that roughly billion-year-old rocks like those characterizing the Grenville province had been found near one Antarctic shore. The Grenville province is an aged band of rocks running along the eastern and southern margin of North America, from Labrador to Texas. He called his hypothesis—the idea that the continents had been juxtaposed—SWEAT, for Southwest U.S.- East Antarctica.

Fired up by the possibility that my question might finally have an answer, I reproduced Moores's reconstruction using the PLATES software at our institute at the University of Texas at Austin. The program allows us to group together pieces of continents and move them over the globe with geometric precision. A short time later my colleague Lisa M. Gahagan and I had removed any uncertainties about matching the boundaries: the scale and general shape of the two old rifted margins were indeed compatible. Moreover, the boundary be-



NORTH AMERICA rifted apart from within the supercontinent Rodinia (*a*), traveled around neighboring continents and rejoined them to form Pangea (*g*). On its way it collided with the South American shield (*d* and *f*) and northern Europe (*e*), leaving characteristic fragments. The central figure is a composite that juxtaposes the various positions of North America through its 500-million-year odyssey; the remaining continents are portrayed in their positions of 260 million years ago. The ancient latitudes of these landmasses are reconstructed using magnetic data.

IAN WORPOLE



a 750 MILLION YEARS AGO

b 750 TO 550 MILLION YEARS AGO

c 540 TO 500 MILLION YEARS AGO

tween the Grenville rocks of Texas and the older rocks of Arizona and New Mexico projected into Antarctica—just where I knew there was a similar boundary under the ice, between the Shackleton Range and some tiny rock outcrops along the frozen shores of the Weddell Sea. It seemed as if the rocks right under my feet, those that form the Llano uplift in Texas and from which the Texas State Capitol was built, were reappearing electronically in Antarctica!

If the western edge of North America was joined to East Antarctica and Australia, then some other continent must have rifted off the Appalachian margin. Paul F. Hoffman, now at Harvard University, and I have suggested that the eastern side of the Laurentian shield was wedged against the Precambrian shields of South America, known as Amazonia and Rio de la Plata. In manipulating the three shields on the computer screen, it occurred to me that the Labrador-Greenland prominence of Laurentia might have originated within the recess in the South American margin between Chile and southern Peru, often referred to as the Arica embayment. Both the promontory and the embayment are believed to date from late Precambrian times. But while they are of the same size and general shape, they were extensively modified when the Appalachian and Andean mountain chains rose. So a precise geometric fit is not to be expected.

My suggestion provides a possible explanation for a long-standing enigma of Andean geology. Along the otherwise youthful and active Peruvian margin are found 1.9-billion-year-old crystalline rocks. Hardolph A. Wasteneys of the Royal Ontario Museum has dated zircon crystals from the Arequipa massif, along the coast of southern Peru. He demonstrated that these rocks were highly metamorphosed when the Grenville Mountains were formed, 1.3 to 0.9 billion years ago. They may therefore represent a continuation of the Grenville province into South America.

The hypothesis of a South American connection for the eastern margin of Laurentia unexpectedly brought my career full circle. I grew up in Scotland and cut my geological teeth on its rocks. Northwestern Scotland and the submerged Rockall Plateau—off the western margin of the British Isles—remained part of North America until the North Atlantic Ocean basin had almost finished opening. Scotland was at the apex of the Labrador-Greenland promontory. When nestled (electronically) in the Arica embayment, the rocks of the Scottish Highlands that I studied for my doctoral degree in the 1960s appear to continue into equally old rocks of Peru and Bolivia. Given how well studied the Scottish Highlands are, they may provide critical tests for a former North America–South America connection.

Assuming the SWEAT hypothesis and the Pan-American connection, we can try to reconstruct the global distribution of continents and oceans in the late Precambrian. Most geologists believe that the relative areas occupied by continents and ocean basins have not changed since the late Precambrian. If, therefore, Antarctica, Australia, North America and fragments of South America were fused into a supercontinent, now named Rodinia, then there had to have been vast oceans elsewhere. Ophiolitic relics caught up within the continents indicate that these oceans lay between India and today's East Africa (the Mozambique Ocean) and within Africa and South America (the Pan-African and Braziliano oceans, respectively).

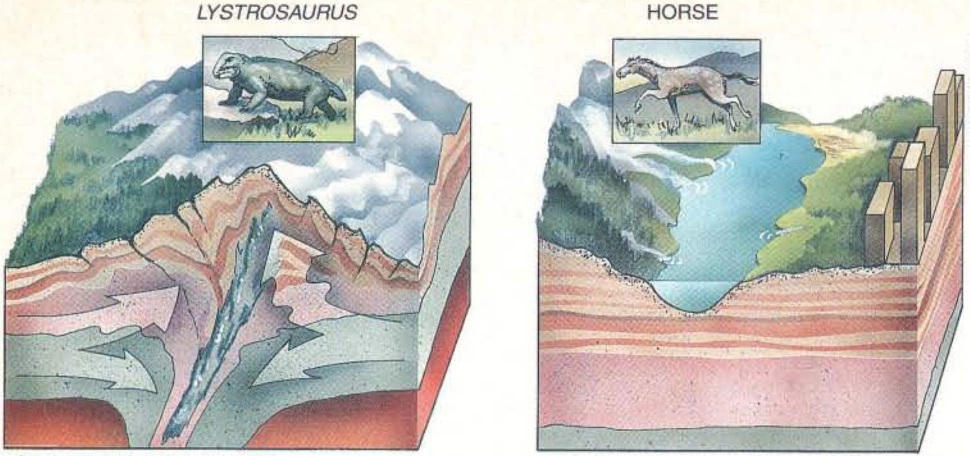
Between roughly 750 and 550 million years ago these ocean basins were destroyed, and all the Precambrian nuclei of Africa, Australia, Antarctica, South America and India amalgamated into the supercontinent of Gondwana. It was during this time interval that the Pacific Ocean basin opened between Laurentia and the East Antarctic–Australian landmass. Isotopic dating of volcanic rocks in Newfoundland shows that the ocean basin between Laurentia and South America did not open until the beginning of the Cambrian. North America may therefore have separated out in a two-stage process.

Calling Cards

Reconstructing the travels of North America requires an essential piece of information: the magnetization of ancient rocks. Such data allow geologists to figure out the latitude and orientation of the rocks when they formed. But because the earth's magnetic field is axially symmetrical, paleomagnetic measurements cannot tell us about the original longitude of the rocks. Present-day lava from Iceland and Hawaii, for example, could reveal to a geologist 100 million years from now the latitudes and the orientation of these islands but not their vast difference in longitude. It would not be apparent that the islands are in different oceans.

Traditional reconstructions of Laurentia always place its Appalachian margin opposite northwestern Africa during the Paleozoic era. I decided to plot the relation of North America to Gondwana differently, taking advantage of the fact that the longitude of the continent is not constrained by paleomagnetic data. It turned out that North America could have made what one of my graduate students referred to as an "end run" around South America during the Paleozoic, starting from next to Antarctica.

When Luis H. Dalla Salda, Carlos A. Cingolani and Ricardo Varela of the University of La Plata in Argentina saw the sketch of the "end run," they became excited. They had recently proposed that a Paleozoic mountain belt, whose roots are exposed in the Andes of northern Argentina, could have formed when another continent collided with Gondwa-



d 250 MILLION YEARS AGO

e 180 MILLION YEARS AGO TO PRESENT

JANA BRENNING

CONTINENTAL DRIFT is reflected in geology characteristic of each era. As Rodinia fragmented (*a*), sandstones of the Patuxent Formation in Antarctica were deposited. Glaciers lined the uplifted shoulders of the rift. In late Precambrian times (*b*), the Pacific Ocean opened up by spreading of the seafloor. Soft-bodied fauna simultaneously developed. Oceans advanced over the continents in Cambrian times (*c*), inhabited by hard-bodied trilobites and reef-forming Archaeocyatha. By the time tetrapods roamed, the continents had amalgamated into Pangea (*d*). The Atlantic and Indian oceans opened as Pangea broke apart (*e*) into today's world.

na. Moreover, the western margin of this Famatinian belt includes Cambrian and Lower Ordovician limestones (between 545 and 490 million years old) containing trilobites characteristic of North America. Perhaps, they reasoned, this is a "geologic calling card" left behind when North America collided with South America during the Ordovician period, 450 million years ago.

It appears that after rifting from South America at the end of the Precambrian, North America moved quite far away. During the Cambrian period, when Gondwana was undergoing glaciation, North America was equatorial. Ocean floor was then subducted beneath the South American craton, and North and South America collided again during the Ordovician. We think that the older part of the Appalachian Mountains, which terminates abruptly in Georgia, was once continuous with the Famatinian belt. This construction places Washington, D.C., close to Lima, Peru, during mid-Ordovician times.

End of the Run

After the collision, the continents separated again, apparently leaving North American limestone with its characteristic trilobites in northwestern Argentina. My Argentine colleagues and I have suggested that these rocks tore off the ancestral Gulf of Mexico, known as the Ouachita embayment. Blocks carried up by Andean volcanoes from below the limestones have recently been dated at around one billion years old, just like those of the Grenville province that probably occupied the embayment.

It is possible that the North and South American continents interacted again before North America finally collided with northwestern Africa to complete Pangea. French geologists studying the Paleozoic sedimentary rocks of the Peruvian Andes have found that they are made of debris that must have eroded from a neighboring landmass. They assumed this continent, occupying the area now covered by the Pacific Ocean, to have been an extension of the Arequipa massif in Peru.

It may, however, have been North America. As Heinrich Bahlburg of the University of Heidelberg has pointed out, ancient warm-water North American fauna mingle with cold-water fauna of southern Africa and the Falkland (or Malvinas) Islands in the 400-million-year-old (Devonian) strata of northwestern South America. Together with a deformation along the eastern seaboard of North America known as the Acadian orogeny, and the truncation of mountain structures along the South American margin, the evidence points to Laurentia's sideswiping northwestern South America during the Devonian. There are even Ordovician limestones with South American trilobites—another calling card—at Oaxaca in Mexico. Only after North America finally moved away from the proto-Andean margin did the Andean Cordillera of the present day begin to develop.

Some 150 million years later North America returned to collide with northern Europe, Asia and Gondwana. Pangea—with the Urals, the Armorican Mountains of Belgium and northern France, the Ouachitas and the youngest Appalachians as sutures—arose from the collisions of these continents. After a 500-million-year odyssey, North America had finally found a resting place. But not for long. In another 75 million years it separated from Africa as Pangea broke up, to move toward its current position.

During the southern summer of 1993–94—six years after my first glimpse of the Pensacola Mountains and glimmerings of North America's odyssey—I returned to Antarctica. This time, with my colleague Mark A. Helper, two graduate students and two mountaineers, I explored the Shackleton Range and Coats Land, near the Weddell Sea. According to my computer simulations, this is where North America's Grenville rocks had projected 750 million years ago. Antarctic geologists have long regarded these areas as anomalous.

At the end of our visit to Littlewood Nunataks in Coats Land, we roped together, picked up our ice axes and climbed back to another small aircraft. Weighing down our packs—and the aircraft, which groaned into the air—were our gleanings of the day. The first paleomagnetic data from those samples seem to support the SWEAT hypothesis. In the laboratories of my colleagues Wulf A. Gose and James N. Connelly, we are still busily taking these rocks apart.

FURTHER READING

SOUTHWEST U.S.–EAST ANTARCTIC (SWEAT) CONNECTION: A HYPOTHESIS. E. M. Moores in *Geology,* Vol. 19, No. 5, pages 425–428; May 1991.

DID THE BREAKOUT OF LAURENTIA TURN GONDWANALAND INSIDE-OUT? Paul F. Hoffman in *Science,* Vol. 252, pages 1409–1412; June 7, 1991.

PACIFIC MARGINS OF LAURENTIA AND EAST ANTARCTICA–AUSTRALIA AS A CONJUGATE RIFT PAIR: EVIDENCE AND IMPLICATIONS FOR AN EOCAMBRIAN SUPERCONTINENT. I.W.D. Dalziel in *Geology,* Vol. 19, No. 6, pages 598–601; June 1991.

ON THE ORGANIZATION OF AMERICAN PLATES IN THE NEOPROTEROZOIC AND THE BREAKOUT OF LAURENTIA. I.W.D. Dalziel in *GSA Today,* Vol. 2, No. 11, pages 237–241; November 1992.

PALEOZOIC LAURENTIA-GONDWANA INTERACTION AND THE ORIGIN OF THE APPALACHIAN-ANDEAN MOUNTAIN SYSTEM. I.W.D. Dalziel, Luis H. Dalla Salda and Lisa M. Gahagan in *Geological Society of America Bulletin,* Vol. 106, No. 2, pages 243–252; February 1994.

Elastic Biomolecular Machines

Synthetic chains of amino acids, patterned after those in connective tissue, can transform heat and chemical energy into motion

by Dan W. Urry

The motion of living organisms is an enigmatic phenomenon. It seems to come spontaneously from within, in the blink of an eye, in the twitch of a muscle, in the sudden kick of a runner, in the jerk-and-press of a weight lifter. Theirs is not the motion of the jet airplane, the rocket or the truck, which results from the explosive expansion of high-temperature gases. Nor is it externally imposed, as in the motion of sailing ships, waves or trees blowing in the wind.

Every living organism represents the successful integration of many biomolecular machines that convert energy from light or raw chemical form into whatever the organism needs—motion, heat or the construction and disposal of internal structures. During the past four decades, biochemists have reverse-engineered the design of the proteins that form the principal components of these machines. Most recently my colleagues and I have progressed to designing artificial models that can carry out the energy conversions that occur in living organisms. These elastic molecules can stretch or contract in response to chemical and electrical signals, or they can generate chemical outputs in response to mechanical stimulation. In theory, these molecules can convert any stimulus into any other form of energy.

Synthetic biomolecular machines can mimic the functions of particular proteins, but they are much simpler. Because they pare away nonessential elements, they can lead biologists to a better understanding of basic cellular mechanisms. Indeed, I believe that the principle underlying their operation may be responsible for the entire range of biological energy conversions.

These molecules also have a broad potential for medical and nonmedical use. They can deliver drugs, prevent the formation of scar tissue or wick liquids away from skin. Their further development could portend a new kind of servomechanism that combines sensors for heat, light, pressure or chemical change with the capacity for motion.

DAN W. URRY has been studying synthetic protein chains for more than 20 years. He is professor of biochemistry, physiology and biophysics at the University of Alabama at Birmingham, where since 1970 he has directed the Laboratory of Molecular Biophysics. He received his B.A. in medical biology and his Ph.D. in physical chemistry from the University of Utah.

Heat Increases Order

A quarter of a century ago my colleagues and I began building on the work of the late Miles Partridge of the University of Bristol in England, who first analyzed the elastic structural proteins of connective tissue. We purified elastin, a protein found in the arteries, ligaments, lungs and skin of all mammals, and found that a solution containing a particular fragment, alpha elastin, underwent a transition from a disordered state to a more ordered one as we raised the temperature. Between 15 and 25 degrees Celsius, the molecules coalesced into a dense, sticky phase.

All proteins are long chains of linked amino acids, the basic building blocks of life (which in turn consist of carbon-nitrogen backbones distinguished by the side chains attached to them). In 1973 William R. Gray and Lawrence B. Sandberg and their co-workers at the University of Utah discovered that elastin contained repeating amino acid sequences, among them valine-proline-glycine-glycine (VPGG) and valine-proline-glycine-valine-glycine (VPGVG). Since then, we have synthesized polymers consisting of these and other repeating units and studied their structure. Substituting other amino acid residues, such as glutamic acid (E), lysine (K) or isoleucine (I), for one of the valine residues yields related fragments (VPGEG, VPGKG or IPGVG) with somewhat different properties. Incorporating these fragments into a chain can modify the chain's behavior.

We found that, just like their natural exemplars, the synthetic polymer chains start in an extended state and then fold into a helix as the temperature rises [*see illustration at top of pages 46 and 47*]. The helices then aggregate to form filaments whose proportions are the same as those of purified natural elastin, as so elegantly measured by the late Laurie Gotte and his co-workers at the University of Padua.

When we cross-linked the molecules into sheets and bands (using gamma radiation from cobalt 60), they behaved much like similarly shaped pieces of rubber. In particular, the elasticity they displayed was mostly what physical chemists call entropic. The limited elasticity seen in a thin metal wire comes from straining the interatomic bonds that hold it together—when the deforming force disappears, these bonds pull the material back toward its original length. Eventually, however, too much stretching will break the bonds. In contrast, when a rubberlike protein is stretched, the bonds between its backbone atoms remain unstrained. Instead the torsional motion of other atoms around these backbone bonds becomes restricted (and thus the entire molecule becomes more ordered). When the deforming force goes away, entropy—the natural slide toward disorder—restores the molecules to their original, freer configuration.

All of us carry proof that entropic elasticity is resilient. The elastic fibers in human arteries—especially in the aortic arch—typically survive for more than 60 years, undergoing billions of cycles of stretching and relaxation.

The entropic nature of these elastic polymers also points to a basic conundrum in their behavior: How can their structure become more ordered as the temperature increases? Most materials become less orderly at higher temperatures, for example, by melting or vaporizing. For that reason, we refer to the change in the polymers as an inverse temperature transition.

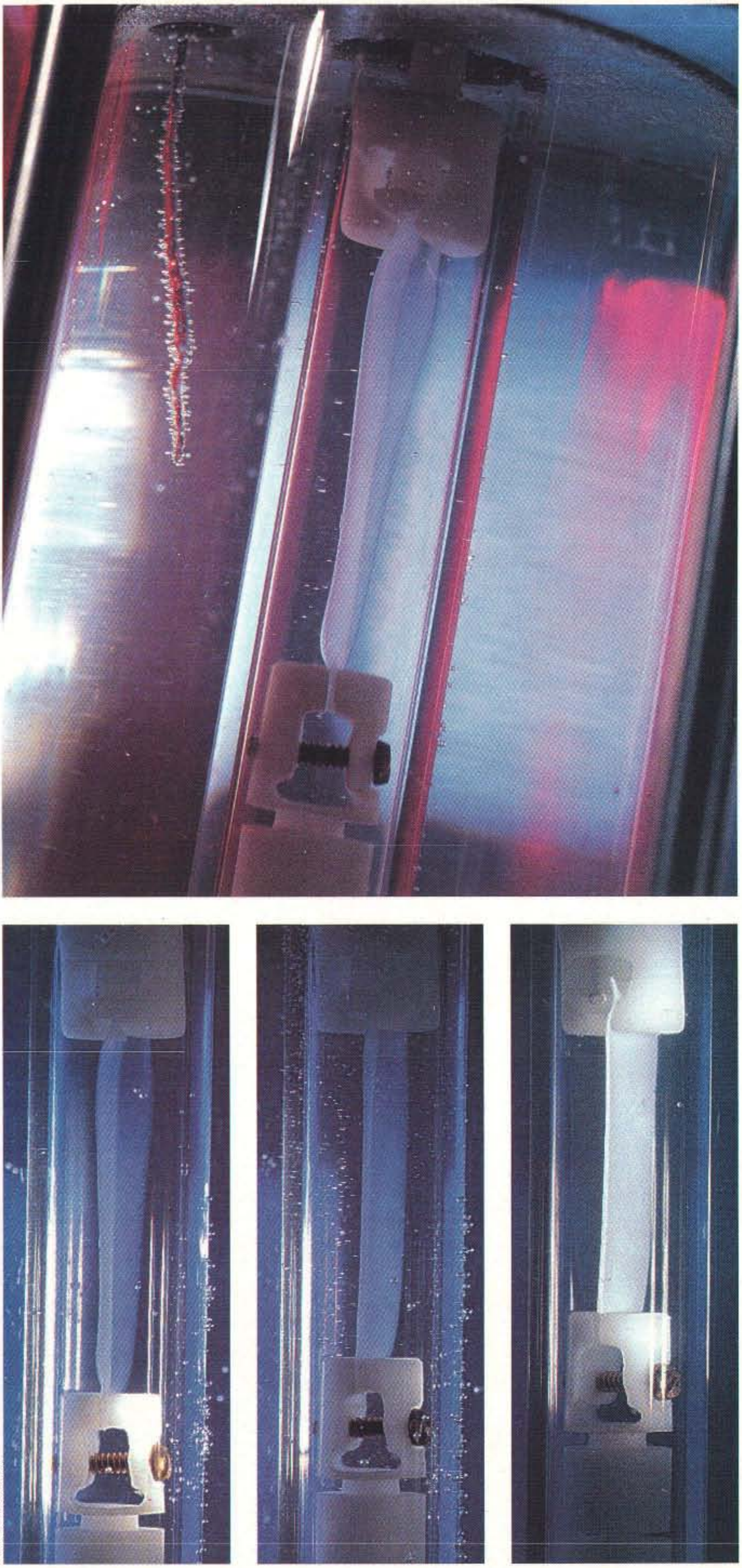
The explanation for this phenomenon, we believe, is in the water that permeates the polymers. Parts of the polymer chain are hydrophobic ("water-fearing," or apolar), and others are hydrophilic ("water-loving," or polar); water molecules accordingly arrange themselves around these sections of the molecules in different configurations. The relative stability of those configurations changes with temperature, and so does the preferred shape of the protein.

Because the hydrophobic segments do not interact strongly with water, the water molecules near them form hydrogen bonds among themselves, creating a series of pentagonal structures. This arrangement of the water helps to keep the polymer in an extended shape: folding would require breaking some of the hydrogen bonds.

At low temperatures the ordered pentagonal structure is stable. With increasing temperature, however, comes a tendency to disorder, and at some point the random thermal energy of the water molecules suffices to disrupt the hydrogen bonds. The pentagonal arrangement of water molecules then gives way to the amorphous, shifting bonds that characterize liquid water in bulk. Unhindered by the water surrounding it, the protein can assume its folded shape. As required by the second law of thermodynamics, the increase in the disorder of the water is greater than the increase in the order of the chain: the total entropy of the system rises.

Altering the Transition Temperature

To put this mechanism to work, we placed our elastic sheets and bands in a water bath and raised the temperature through the range at which we would expect the inverse temperature transition to occur. As we had expected, the materials contracted to one half their swollen length. By moving in response to a temperature change, these biopolymers had converted heat into work, just as conventional engines do. This conversion turned out to be highly

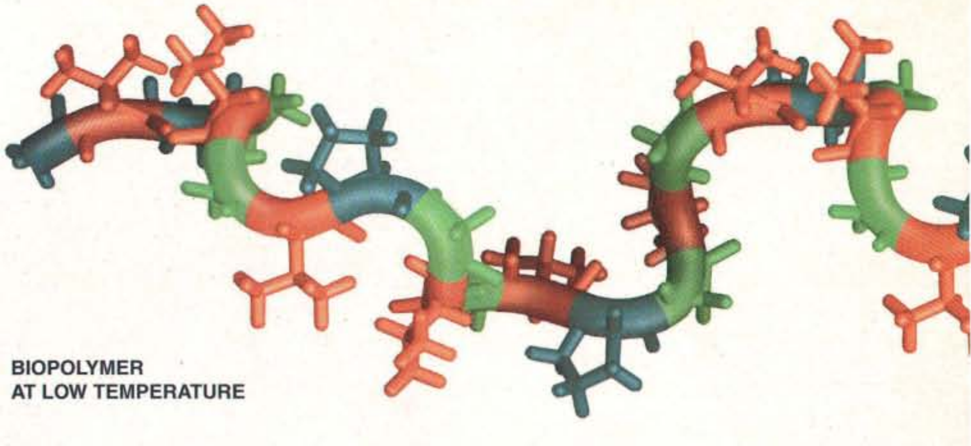


SYNTHETIC PROTEIN SHEET (*between clamps, top*) lifts a small weight when it is heated, converting thermal energy to work (*bottom*). The sheet shrinks and becomes less transparent as its internal order increases. It consists of cross-linked chains of amino acids; each chain contains 100 or more repetitions of the sequence valine-proline-glycine-valine-glycine.

ANN STATES SABA

efficient: some biopolymers can lift more than 1,000 times their dry weight.

The precise temperature at which the transition takes place depends on the mix of polar and apolar segments in the polymer. Hydrophilic side chains on the amino acids tend to raise the transition temperature; hydrophobic side chains lower it. The reason for this trend is that hydrophilic groups such as carboxylate (COO^-) exert a strong pull on water molecules and so can destroy the pentagonal-structure water surrounding adjacent hydrophobic side chains. Consequently, a temperature that would disrupt the order of water around a series of hydrophobic groups will have less effect when hydrophilic chains are nearby, because the water has already been disordered.



Adding hydrophilic groups to a chain thus increases the transition temperature. At the same time, however, the hydrophilic groups reduce the amount of energy required for folding because only a few remaining pentagonal water bonds need be disrupted. In one biopolymer, inserting just two units of carboxylate per 100 backbone amino acid residues can reduce the amount of heat required to fold it by 75 percent.

Of course, when animals move, their muscles respond to electrochemical signals, not to temperature changes. To mimic this behavior, we have made variants of our original bioelastic polymer in which a stimulus can trigger folding or unfolding by altering the transition temperature. The thermal environment remains constant, but the molecular chain's response to it changes. (I call this effect the ΔT_t- mechanism because of the way it manipulates the transition temperature, T_t.)

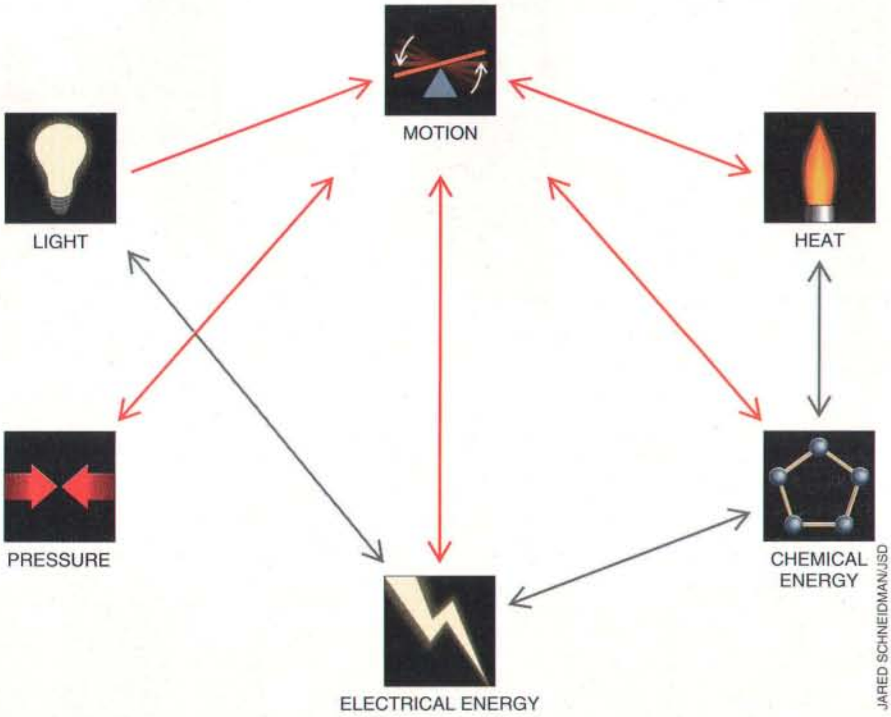
The key to this legerdemain is an amino acid with a side chain that can change its behavior from water-fearing to water-loving (and back again), depending on conditions around it. Our first success came with glutamic acid, whose side chain takes the form of COOH under acid conditions and COO^- at neutral pH. COOH is more hydrophobic than is the charged COO^-, and so at body temperature a band containing glutamic acid should contract at low pH and swell at the roughly neutral pH prevalent inside a living organism.

Theory into Practice

We demonstrated this behavior with a sheet made of poly[4(VPGVG), (VPGEG)], a polymer in which one of every five repeating amino acid sequences contains a glutamic acid residue in place of a valine. When we changed the pH from seven to four, the sheet contracted and lifted a weight. To our knowledge, this experiment marked the first time that a synthetic model protein had been made to convert chemical energy reversibly into motion.

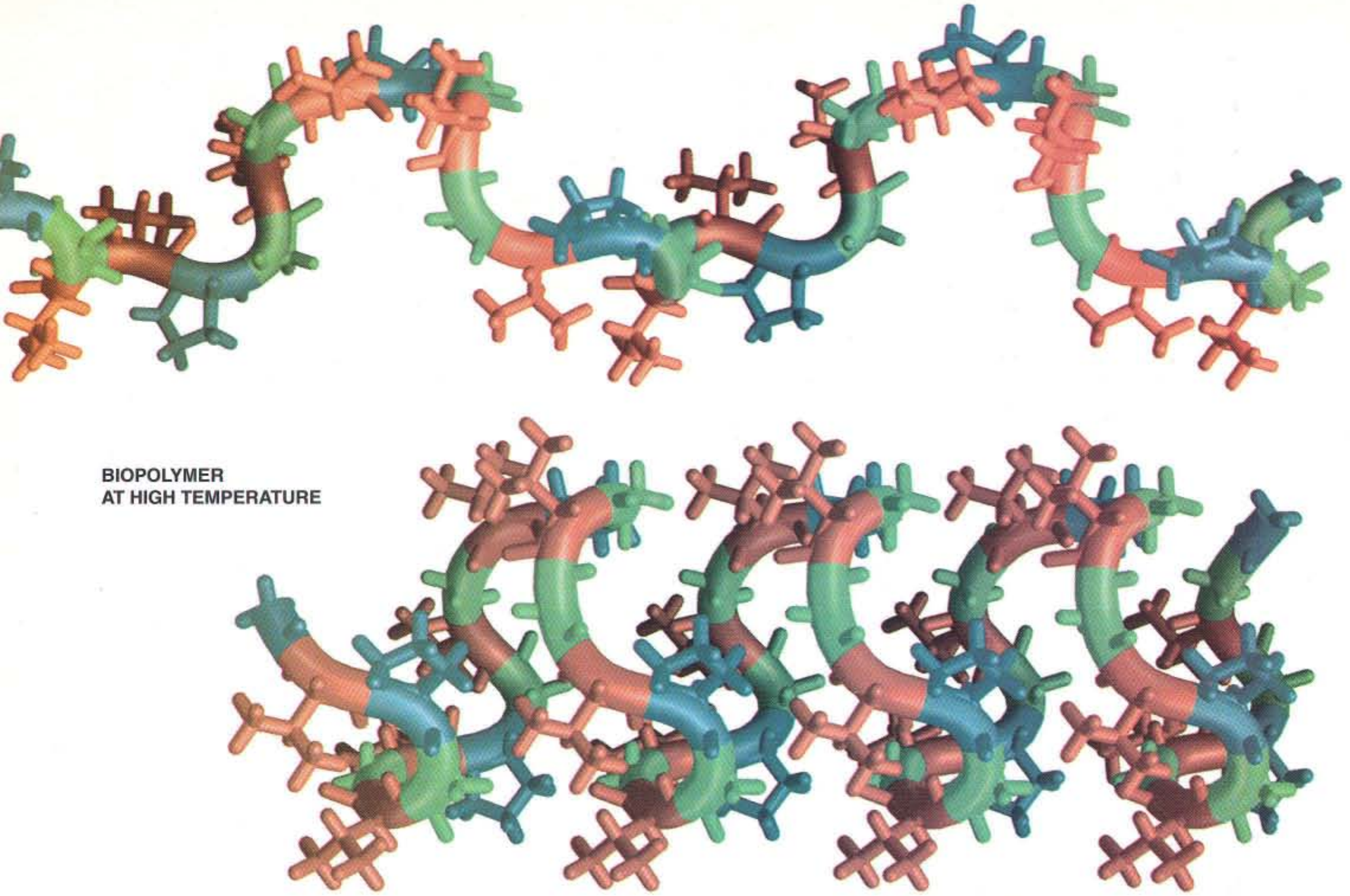
As with the purely temperature-driven transition, the chemical conditions under which folding takes place depend on the mix of hydrophilic and hydrophobic side chains. If a polymer's side chains are mostly hydrophobic, a great deal of pentagonal water must be destructured before the chain can contract. The energy requirement can shift the transition to a higher pH. For example, if the glutamic acid is surrounded by several very hydrophobic groups, its side chain will remain in the COOH state at typical body pH. Yet the transition, once initiated, will go to completion over a much smaller pH range. All the COOH groups will change to COO^- when the pH changes from seven to eight (instead of the more usual range from four to seven).

We have since demonstrated that we can make a protein chain expand or contract in response to any reversible change that raises or lowers its transition temperature. For example, we can convert electrical energy to motion by attaching a so-called prosthetic group (a molecular fragment that readily accepts or gives up electrons) to the chain. The prosthetic group is more hydrophobic in its reduced state, after it has gained an electron, than in its oxidized state. As a result, adding electrons will lower the chain's transition temperature, thereby causing it to fold.



JARED SCHNEIDMAN/JSD

ENERGY can be converted between many different forms by biomolecular machines. Some conversions (such as heat to motion) proceed directly (*red arrows*), whereas others require an intermediate step (*gray arrows*).



ANDREAS WINDEMUTH *using GRASP by Anthony Nicholls and Barry Honig, Columbia University*

AMINO ACID CHAIN stretches at low temperatures (*top*) and folds at higher ones (*bottom*). The amino acid residues valine (*blue*), proline (*orange*) and glycine (*green*) bond to one another to form one third of a spiral turn. Repetitions of this basic unit give rise to a helical structure, about 1.8 nanometers across, known as a beta spiral. The folding or extension of the spiral depends on interactions between water and hydrophobic side chains on some of the amino acids.

A biopolymer that contains hydrophobic residues such as phenylalanine can convert changes in pressure to motion. The chains fold at low pressure and extend when higher pressures favor the pentagonal water structure. Finally, we can convert light to motion by attaching a photosensitive group: either azobenzene or cinnamide will switch from a more hydrophobic to a more hydrophilic state when struck by light of the appropriate wavelength.

A Universal Mechanism?

As biochemists take apart living cells at the molecular level or synthesize exact duplicates of their components, it will be possible to determine whether the inverse temperature-transition mechanism is in fact the one that living organisms use. All the available evidence points in that direction. First, it is the most efficient mechanism known for driving the folding and unfolding of proteins. Natural selection ought by rights to have found it over the course of three billion years. Indeed, it is the most efficient method known in polymers for converting chemical energy into motion. Polymers that expand or contract in response to repulsion between charged subunits are not nearly so effective, because the surrounding water molecules shield the charges from one another. Polymethacrylic acid, for example, must convert 15 times as many residues from charged to uncharged to accomplish the same work as one of our biopolymers.

Second, phosphates—the energy currency of choice in most organisms, in the form of ATP (adenosine triphosphate) and its equivalents—are extraordinarily good at driving changes in the transition temperature. Adding a phosphate is four times as effective as converting carboxylic acid to carboxylate. Bound ATP, which acts as a multiply charged phosphate group for these purposes, is equally potent.

Testing the effect of phosphates on our molecular chains was not simple. We had to create a polymer containing a sequence of amino acids that could be phosphorylated or dephosphorylated easily by adding enzymes to the fluid surrounding it. Success came with a subunit known as RGYSLG (arginine-glycine-tyrosine-serine-leucine-glycine) and an enzyme called a protein kinase, which adds phosphate to the serine residue. By inserting a 3 percent fraction of this sequence in our chains, we made polymers that would unfold in the presence of the protein kinase. Alkaline phosphatase, which removes the phosphate, caused the chains to fold again.

The phosphate-driven mechanism that we demonstrated in our biopolymer could also control the folding and assembly of proteins inside a cell. A particular protein, for example, might fold into a closed position at body temperature because its environment is too hot for pentagonally bonded water to form around its hydrophobic surfaces. A single phosphate or a molecule of ATP bound to the protein would permit water to surround the area in destructured form and so shift the equilibrium toward the open position.

By forcing such a protein open, the phosphate could expose bonding sites that would otherwise be hidden from other proteins, thus facilitating crucial interactions. For example, "heat shock" proteins (also known as chaperonins) help some proteins with significant hydrophobic surfaces to fold properly. Interestingly, one such chaperonin has a form homologous to the component of muscle that binds ATP.

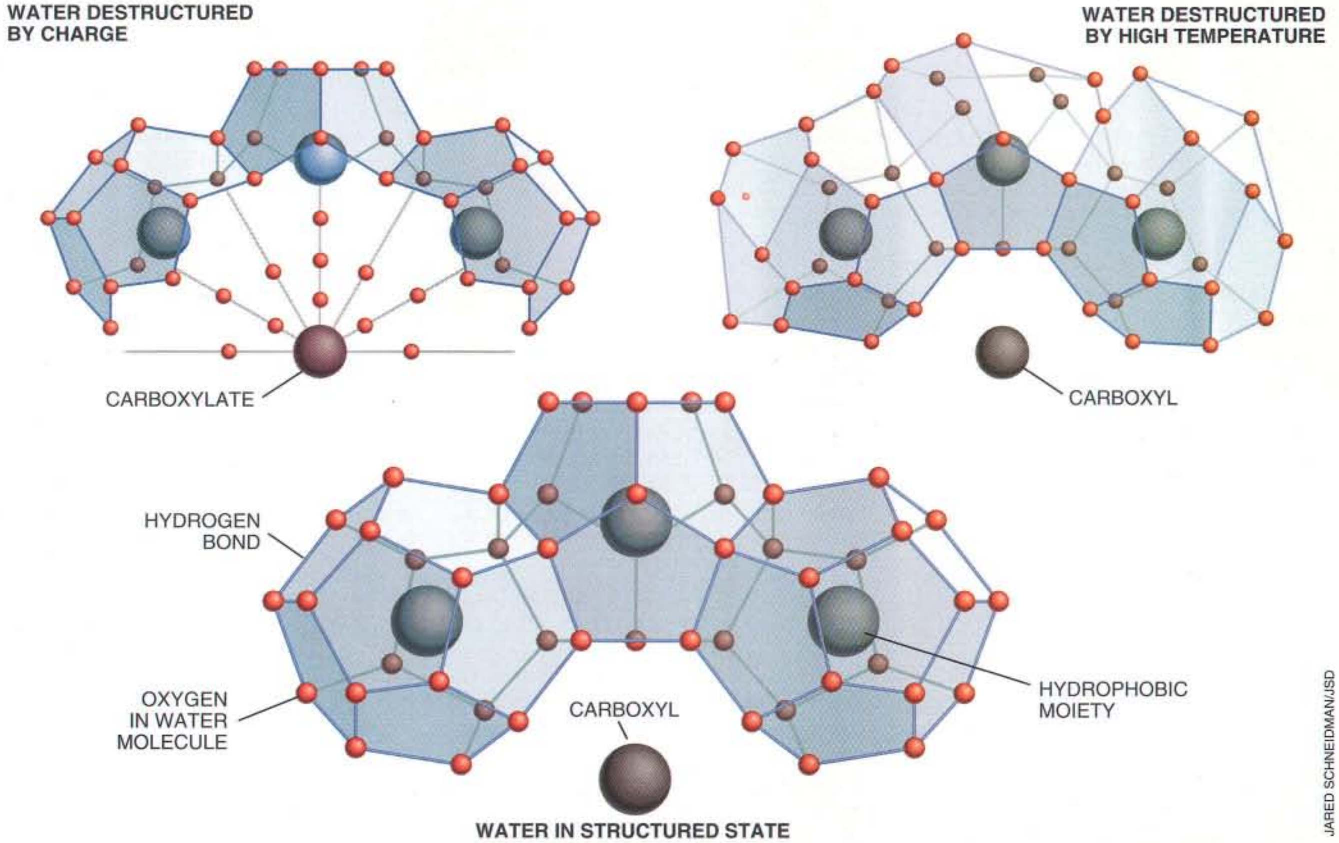
I believe that the inverse temperature-transition mechanism is at the heart of most biological conversions of energy. Oxidative phosphorylation, the process whereby the chemical energy in food is converted to ATP, can be understood largely in terms of this mechanism. Inside cells the oxidation and reduction of prosthetic groups produce gradients of protons, which in turn drive the phosphorylation reactions that form ATP. All these energy-conversion steps could be accomplished by proteins' folding or unfolding in response to changes in their transition temperatures.

Building biopolymers that react to ATP and its associated enzymes will help researchers understand many basic cellular mechanisms. Such knowledge could eventually lead to synthetic replacements for muscles and a variety of tissues. Bioelastic materials are particularly promising as tissue substitutes because they can be made in a range of consistencies. Depending on the amino acids in the polymer chains, the result may be gelatinous, rubbery or as rigid as plastic. Matching the elastic properties of the tissue being replaced, either temporarily or permanently, is crucial to integrating the synthetic with the living tissue around it. Cells can sense the tension to which tissue is subjected; stretching (or compression) appears to trigger chemical responses that cause the cell to function appropriately for its particular location.

Long before bioelastic polymers become satisfactory tissue substitutes, however, they will have a host of other applications in the body. Unlike most other materials, elastic sheets of poly(VPGVG) implanted in rabbits and rats appear to be simply ignored by the immune system. The most immediate application of this biocompatibility could be in preventing adhesions, the painful and potentially dangerous overgrowths of tissue that often occur after surgery or trauma. A sheet of polymer acting, in essence, as an internal bandage can keep adjacent layers of tissue from fusing together.

Lynne D. Hoban, Taffy Williams and Adam McKee, working at the Naval Medical Research Institute, placed a sterilized sheet of poly(VPGVG) in the abdominal cavities of rats during surgery that usually causes adhesions. They found that the material reduced destructive tissue growth to an insignificant level in 80 percent of their subjects. Half of the adhesions that did form occurred because the sheet had been torn during the operation. In the remaining cases, tissue overgrowth engulfed the entire sheet, but the polymer separated readily from the adhesion.

These materials also look promising in preliminary studies of eye surgery. Surgeons correct strabismus (a misalignment of the eyes) by repositioning


WATER DESTRUCTURED BY CHARGE
WATER DESTRUCTURED BY HIGH TEMPERATURE
CARBOXYLATE
CARBOXYL
HYDROGEN BOND
OXYGEN IN WATER MOLECULE
CARBOXYL
HYDROPHOBIC MOIETY
WATER IN STRUCTURED STATE

JARED SCHNEIDMAN/JSD

PENTAGONAL WATER structure is the result of hydrogen bonds between water molecules adjacent to hydrophobic molecular groups. This low-energy state is stable at low temperatures (*center*) but becomes progressively more unstable at higher ones (*right*). Charged, hydrophilic molecules can destroy this pentagonal structure (*left*) by forcing water molecules to line up around them. Pentagonal water surrounding an amino acid chain can prevent it from folding because energy is required to break the hydrogen bonds so that the water can move out of the way.

COURTESY OF DAN W. URRY

FIBROBLASTS grow on the surface of a biopolymer whose structure has been modified to include chemical sites for cell attachment. (Unmodified versions implanted in the body provoke neither cell growth nor inflammation.) Such polymers could temporarily replace arteries or other tissues.

some of the muscles that control the gaze. Unfortunately, this procedure often leaves scar tissue on the muscle, impairing its function. Frederick J. Elsas of the University of Alabama at Birmingham's Eye Foundation Hospital, operating on rabbits, found that wrapping the cut muscles in a biopolymer sleeve prevented this scarring. These elastic bands also appear to hold promise for surgery to correct detached retinas; they can be wrapped around the eyeball to compress it and promote healing after the retina has been tacked in place by a surgical laser.

Synthetic polymer chains can also incorporate biologically active amino acid sequences. A cross-linked sheet containing one unit of the well-known cell-attachment sequence glycine-arginine-glycine-asparagine-serine-proline (GRGDSP) for every 40 units of GVGVP is still clear and elastic. Instead of ignoring this sheet, however, cells adhere to it, spread and grow to confluence. A temporary functional scaffolding of bioelastic materials for an artificial blood vessel, for example, could be designed so that cells would invade and convert the matrix into natural tissue even while the polymer served its required structural role.

Integrated Drug Delivery

More complex biopolymers might be able to mimic not the mechanical functions of tissues but rather their chemical behavior. A block containing therapeutic agents could release them either at a fixed rate over time or by means of built-in sensing mechanisms—what Ralph Christoffersen of Ribozyme Pharmaceuticals calls a diagnostic-therapeutic pair.

Such sustained-release polymers might rely on chemical clocks that convert an uncharged carboxamide (CONH$_2$) to carboxylate (COO$^-$) over a period that can range from days to decades. In combined diagnostic-therapeutic devices, designers could couple a combination of enzymes and prosthetic groups to the molecular chains so that conditions in the body would control folding or unfolding automatically. A biopolymer to control diabetes, for example, might incorporate a prosthetic group coupled to the enzyme glucose reductase. As blood sugar increased, the enzyme would transfer electrons to nearby glucose molecules from the prosthetic group. Newly hydrophilic, the polymer would swell, releasing insulin.

Elastic biopolymers can be used in more mundane ways as well. Because the conversions that control folding are generally reversible, the molecules can convert mechanical work to chemical or other forms of energy just as easily as they convert light, heat or pressure to motion. A biomolecular desalination plant might rely on a polymer belt that normally shrinks in the presence of salt. The belt would be stretched in briny water, opening up space for water molecules (but not dissolved salts) to form pentagonal structures around its hydrophobic side chains. Rollers in a discharge tank would squeeze out the purified water.

Another, more down-to-earth application arises out of the polymers' combination of water absorption and temperature sensitivity: a uniquely effective biodegradable diaper. Consider a diaper made of a molecular chain that experiences an inverse temperature transition just below the normal temperature of skin. When liquid touched the inside surface, the molecules there would soak it up, but they would remain in the folded state. The polymer chains' propensity to unfold at lower temperatures would spontaneously wick fluid away from the warm body and toward the cool outer surface. The dry matrix could absorb 10 times its own weight in water, and after use the diaper would be readily degraded by microorganisms in the environment. Genetic engineers have already placed the DNA sequences for making several bioelastic materials into *Escherichia coli* bacteria; mass production in fermenting tanks could eventually make the cost of biomaterials competitive with that of the organic polymers now used in disposable diapers.

This simple design is but one of the many nonmedical uses possible for elastic biopolymers. Bioelastic materials have so many potential uses in part because they can be designed to interconvert so many forms of energy. They can sense almost any change in the environment around them, and they can produce chemical, electrical or physical responses. Once these materials can be made in commercial quantities, they should be cheap enough to apply almost anywhere. Implanted molecular clocks, moreover, will permit products made from them to degrade gracefully once their working lives are finished.

Perhaps in the next century biomechanical devices will transform everyday life as thoroughly as electrical devices have in this century. My colleagues and I look forward to a busy future designing elastic biomolecular machines.

FURTHER READING

MEDICAL APPLICATIONS OF BIOELASTIC MATERIALS. D. W. Urry in *Biotechnological Polymers: Medical, Pharmaceutical and Industrial Applications.* Edited by Charles G. Gebelein. Technomic Publishing, 1993.

MOLECULAR MACHINES: HOW MOTION AND OTHER FUNCTIONS OF LIVING ORGANISMS CAN RESULT FROM REVERSIBLE CHEMICAL CHANGES. Dan W. Urry in *Angewandte Chemie, International Edition in English,* Vol. 32, No. 6, pages 819–841; June 1993.

POSTULATES FOR PROTEIN (HYDROPHOBIC) FOLDING AND FUNCTION. Dan W. Urry in *International Journal of Quantum Chemistry: Quantum Biology Symposium, Vol. 21,* pages 3–15; 1994.

The Oldest Old

People in their late nineties or older are often healthier and more robust than those 20 years younger. Traditional views of aging may need rethinking

by Thomas T. Perls

In medical school I was taught that the incidence of chronic, disabling disorders, particularly Alzheimer's disease, increases inexorably with age. I therefore expected that people older than 95 years, often called the oldest old, would be my most debilitated patients. Yet when I became a fellow in geriatrics, I was surprised to find that the oldest old were often the most healthy and agile of the senior people under my care. In fact, the morning I was scheduled to interview a 100-year-old man as part of a research project, he told me I would have to delay my visit. He had seen 19 American presidents take office, and he would be busy that morning voting for the next one.

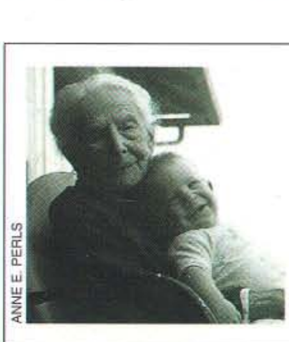
Such encounters made me wonder if the prevailing view of aging as advancing infirmity was partly wrong. Could it be that many people in their upper nineties enjoy good health and that the oldest old constitute a special—and long-misunderstood—population? Since then, the centenarians I have met have, with few exceptions, reported that their nineties were essentially problem free. As nonagenarians, many were employed, sexually active and enjoyed the outdoors and the arts. They basically carried on as if age were not an issue. And accumulating evidence indicates that a significant portion of the oldest old are indeed healthier than many people in their eighties or early nineties. The common idea that advancing age inevitably leads to extreme deterioration does, indeed, seem to require revision.

Estimated costs of caring for the oldest old in the future might need modification as well. The centenarian population grew by 160 percent in the U.S. during the 1980s. Many demographers predict that 20 million to 40 million people will be aged 85 or older in the year 2040, and 500,000 to four million will be centenarians in 2050. The economic burden of caring for people older than 85 could be vast, especially if a huge percentage of them need special care. Yet it may well be that health bills for the oldest old will be lower than previously expected.

Some of the first evidence supporting my suspicions came from a study on Alzheimer's disease that I conducted with my mentor, Lewis A. Lipsitz of the Hebrew Rehabilitation Center for Aged in Boston. Surveys have reported that this disorder devastates the mind and ultimately kills about 40 percent of those aged 85 and older. Some investigators believe that close to 50 percent

JOHN R. McCUTCHEN *San Diego Union-Tribune*

100 YEARS OLD AND STILL SWIMMING, Tom Lane raced in the 100-meter backstroke event of the Senior Olympics, held last September outside San Diego. Lane is among the many healthy centenarians who contradict the traditional idea that age always brings with it severe debilitation. In addition to swimming, Lane also throws the javelin and shot put, bowls and plays golf.



ANNE E. PERLS

THOMAS T. PERLS met his first centenarian in his own family—his great-grandmother, Julia Grunewald, lived to be 102. She is shown with the author at the left. As principal investigator of the New England Centenarian Study, Perls has examined definitions of normal aging and pursued preventive strategies for Alzheimer's disease. He received his medical degree from the University of Rochester and his master's in public health from Harvard University. Perls is an instructor at Harvard Medical School and a geriatrician at Deaconess Hospital in Boston. He is also a recipient of an Alzheimer's Association Faculty Scholar Award.

of 90-year-olds have Alzheimer's disease and that up to 70 percent of centenarians are affected. Yet many of the studies on which these conclusions are based did not include subjects older than 93 years, which casts some doubt on these projections. In 1991 Lipsitz and I undertook a pilot study to see if the occurrence of Alzheimer's disease at the center, a chronic care hospital, matched the predictions for centenarians. We found that of the 12 residents in their hundreds, only four seemed to have Alzheimer's disease. This low figure—only 25 percent—was particularly striking considering that residents of such facilities are more likely to be impaired than are their counterparts in the community.

Selective Survival

Our finding suggested that, at least cognitively, the oldest old were indeed in better shape than has usually been assumed. What, we wondered, could explain their good condition? We suspect that the answer to this riddle is that, for whatever reason, some people are particularly resistant to acquiring the disorders that disable and kill most people before age 90. Because of this resistance, they not only outlive others, they do so relatively free of infirmities. In other words, in a kind of survival-of-the-fittest phenomenon, these individuals seem to be selected for long-term survival because they possess traits that enable them to avoid or delay the diseases that commonly accompany aging.

The concept of selective survival was applied, somewhat more narrowly, by

GEORGES GOBET *Agence France Presse*

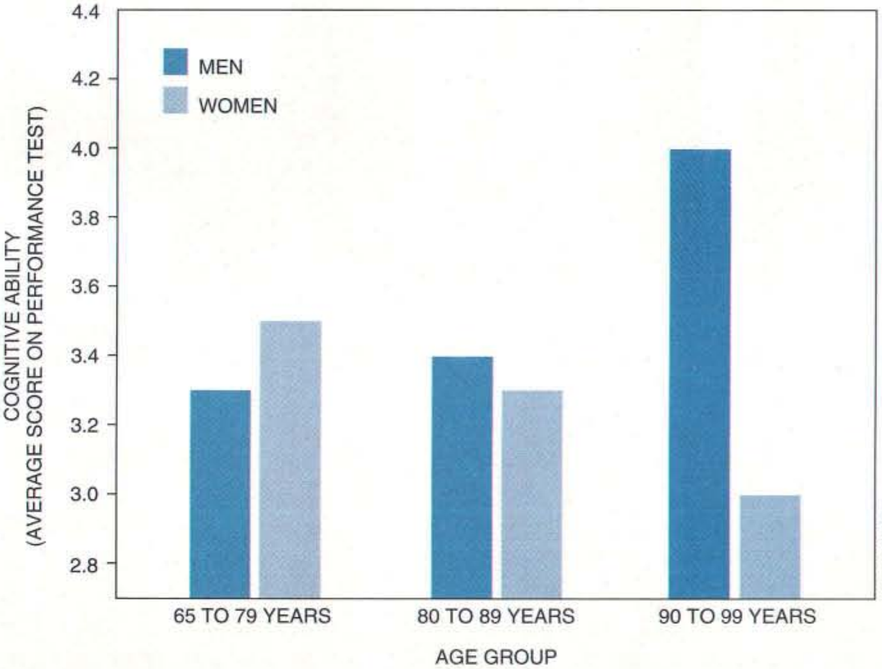
The oldest person

Survival of the Fittest

Jeanne Calment, a Frenchwoman who turns 120 in February, is older than anyone whose age has been confirmed. She is shown at a party celebrating her 116th birthday. Calment is among the people who taught researchers that mortality rates for the oldest old are much lower than would be predicted by extrapolating from the death rates of younger individuals (*left graph*). James W. Vaupel and Anatoli Yashin of Odense University in Denmark, A. Roger Thatcher, formerly of the Office of Population Censuses and Surveys in London, and Vaino Kannisto, formerly of the United Nations, examined death statistics on eight million people. They found that after age 97 a person's chance of dying at a given age veers from the expected trend (*light green*). Instead of increasing exponentially, the rate slows to become more linear (*dark green*). (The ratio would exceed one if an entire age group were to die in less than a year.) These findings support the author's suggestion that the oldest members of our species tend to be healthier than traditional views of aging would predict.

Similar mortality trends were observed among medflies (*right graph*), providing further support for the author's hypothesis. James R. Carey of the University of California at Davis compared expected death rates (*light orange*) with observed rates (*dark orange*). He found that the chance of dying at any given age peaked at around the age of 50 days. After that, the risk began to decline, so that by the age of 100 days, the oldest insects had only a 5 percent chance of dying on a given day.

demographers in the 1970s to older African-American populations. Researchers reported that although the death rates for blacks were higher than for white Americans up to age 75, the trend reversed after that age. Then, in what some called a crossover phenomenon, whites were more likely to die at a given age than were their African-American counterparts. They speculated that blacks tended to die earlier because many of them were economically disadvantaged and had less access to health care services. Therefore, those who survived represented an unusually vigorous group, able to overcome obstacles that defeated others. Their vigor, in turn, later gave them a survival advantage over the majority of white Americans of similar age.



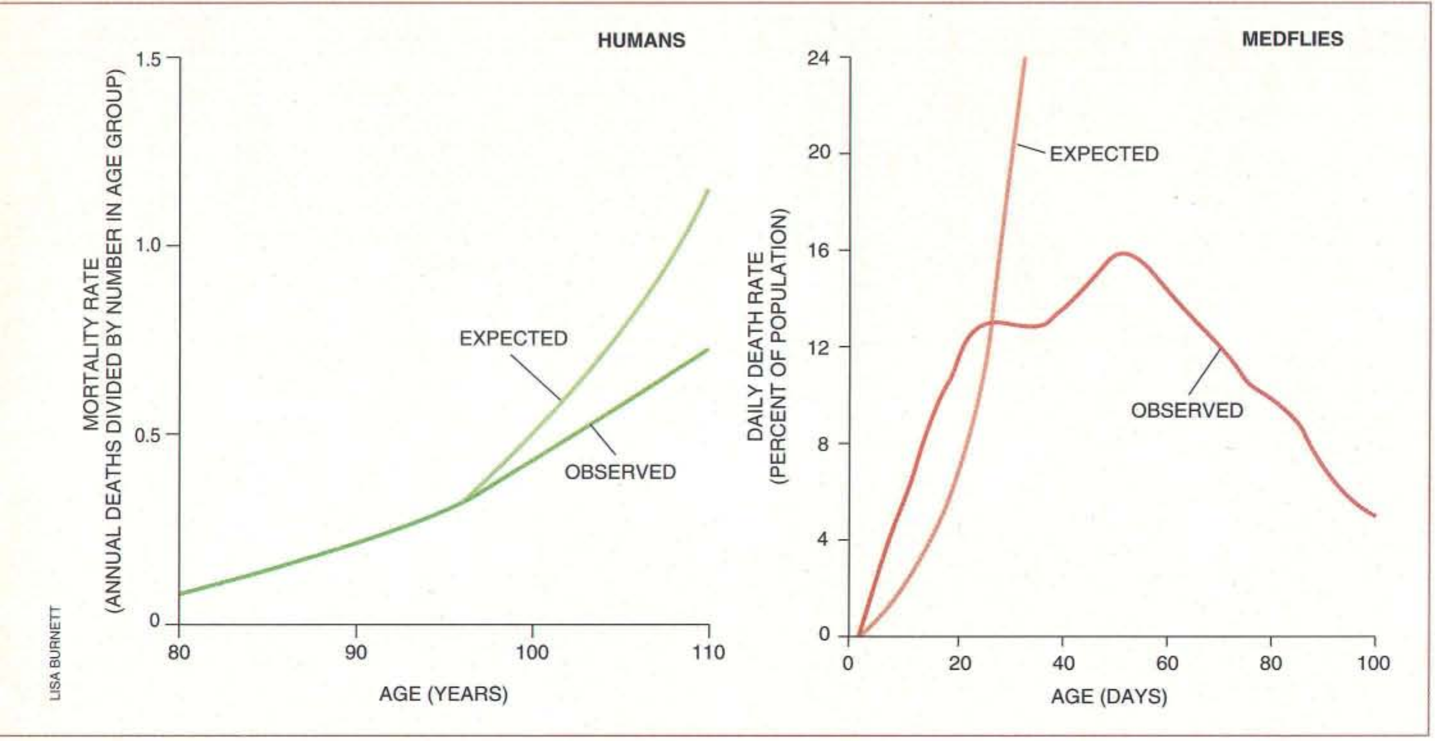
LISA BURNETT

COGNITIVE ABILITIES of oldest old men (*dark blue bar at far right*) are on average higher than the abilities of their female peers (*light blue bar at far right*) even though among people aged 65 to 79 women seem to have a slight advantage (*far left*). The reversal, known as a gender crossover, occurs between the ages of 80 and 89 (*middle bars*). It arises because men who are cognitively impaired generally die earlier than do women, leaving mainly mentally intact men behind.

This selective survival hypothesis may also clarify various other once puzzling findings demonstrating unusually good cognitive and physical health in the oldest old. It seems that men who survive into their late nineties become less and less likely to develop Alzheimer's disease with each passing year. Moreover, the average man in his late nineties has a more intact mind than the average man in his eighties. These patterns probably emerge because men who are susceptible to Alzheimer's disease generally die of the condition in their eighties or early nineties. These trends would be explained if the group of men who reach their late nineties consist almost exclusively of individuals who are not susceptible to Alzheimer's disease and who therefore retain their cognitive abilities indefinitely.

Gender Crossover

Surprisingly, as a group, men older than 90 generally have better mental function than do their female peers. Women, it seems, tend to live with their dementia rather than to die from it. In consequence, very old females on average retain less of their mental abilities than do men of the same age—who represent the healthy survivors left after



LISA BURNETT

other men susceptible to dementia have died off.

At later ages, men also do better than women in terms of physical health. Men in their sixties and seventies are more susceptible than are women to strokes and heart attacks. Delayed onset of these acute conditions allows women to survive longer than men who have to cope with disease earlier in life. In absolute numbers, many more women are still alive at 95, but in terms of average mental and physical health, men begin to take the lead. The healthy men who have generally avoided illnesses demonstrate a survival advantage over women: although men make up 20 percent of 100-year-olds, 40 percent of 105-year-olds are male. This switch to more mentally and physically fit men after age 90 constitutes what I have called a gender crossover.

Early signs of the gender crossover can be seen in studies of 80-year-olds. Men who survive to this age without major health problems often continue to live without needing special care. Richard M. Suzman and his colleagues at the National Institute on Aging found that men older than 80 years in one such study were more independent than were similarly aged women. Their 1984 report indicated that 44 percent of the men in that age group were robust and independent compared with only 28 percent of women in the same age group. Additionally, Kenneth G. Manton and Eric Stallard of Duke University estimated the active life expectancy—that is, the years of independent life left—for members of the U.S. senior population. Their findings showed that after age 85, men could expect to live a healthy and active life longer than women could.

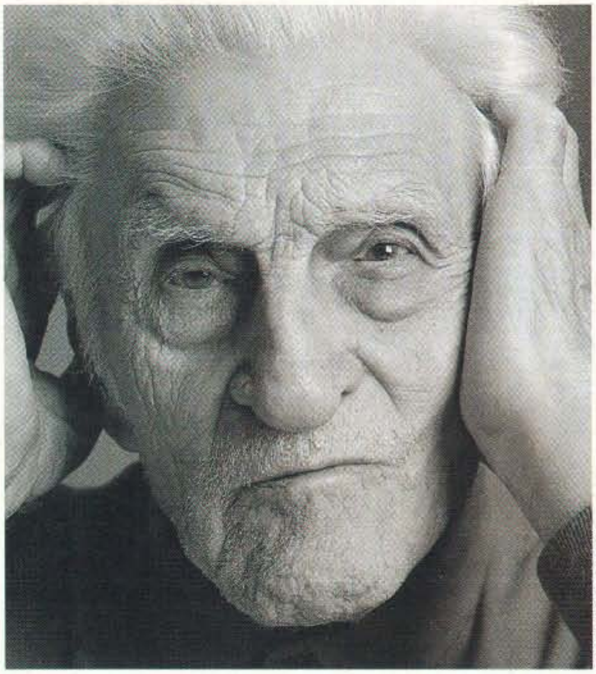
What biological and environmental factors might allow the oldest old humans to reach age 95 and beyond in good health? Multiple and intertwined influences undoubtedly play important roles. So-called longevity genes seem to protect against the development of diseases; genetically or otherwise determined adaptive abilities enable survivors to avoid potentially life-threatening conditions. Modifications in everyday activities, such as not smoking, practicing better nutrition and exercising, may also help some people stay fit longer. Basic good luck surely helps as well.

The Genetic Factor

Tempting candidates for possible longevity genes in humans are ones that control the body's mechanism for protecting itself against oxygen radicals. These naturally occurring, highly reactive compounds damage DNA and can destroy cells. Everyone has a genetically determined ability to combat this type of damage. Gene variants that give rise to unusually efficient resistance to oxidative damage could well contribute to the life span of the oldest old by slowing the rate at which oxygen radicals damage cells.

In addition to carrying longevity genes, the oldest old may have an unusually low complement of deleterious genes. My colleagues and I have found evidence for such a scenario. Recently one variant of the gene coding for the protein apolipoprotein E (*apo-E*) has been tied to a substantially increased risk of acquiring Alzheimer's disease. The average age of onset for Alzheimer's disease appears to be related to the type of *apo-E* genes a person inherits from each parent. There are three common forms: *E2, E3* and *E4*. People who inherit two *E4* genes (one from each parent) have an eight times greater risk than the general population of developing the disease; those with two *E4* genes who acquire the disease display symptoms at an average age of 68. Alzheimer's disease patients with two *E3* genes demonstrate symptoms of the disease somewhat later, at about 75 years. The role of *E2* remains unclear, but there is evidence that it is associated with a lower risk of developing Alzheimer's disease.

In collaboration with Bradley T. Hyman's laboratory at Massachusetts General Hospital, we have determined the prevalence of *E4* among healthy subjects aged 90 to 103. Our study revealed that 14 percent of the group (having an average age of 93) had at least one *E4* gene. Previous studies of 85-year-olds indicated that 18 percent carried at least one *E4* gene, and 25 percent of subjects younger than 65 carried the gene type. The occurrence of the *E4* variant decreases markedly with advancing age, dropping nearly 50 percent over 28



ERIK LEIGH SIMMONS

Three Who Thrive

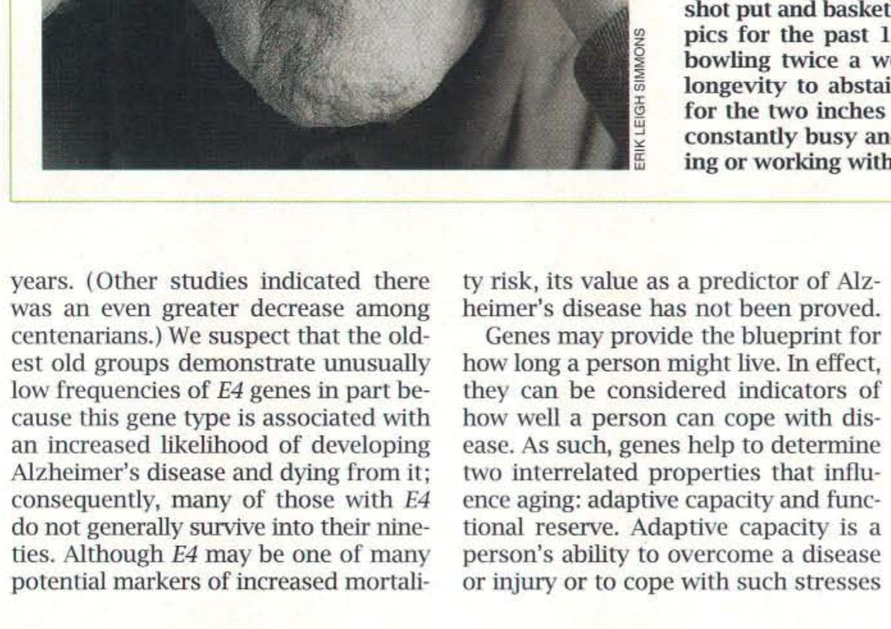
The people shown here are some of the healthiest oldest old who have beaten the odds and survived into their late nineties or beyond. Research suggests that inheriting good genes probably provides the best hope for a long and healthy life. Other factors may be important as well. People such as these have their own hypotheses to explain their longevity; further scientific studies should help clarify the issue.

ALFRED BENEDETTI, age 101, has participated in the javelin, shot put and basketball free-throw events in the Senior Olympics for the past 11 years. Up until last summer, he went bowling twice a week. Benedetti attributes his health and longevity to abstaining from smoking or drinking—except for the two inches of port wine he drinks every day. He is constantly busy and spends much of his time reading, writing or working with his hands.

years. (Other studies indicated there was an even greater decrease among centenarians.) We suspect that the oldest old groups demonstrate unusually low frequencies of *E4* genes in part because this gene type is associated with an increased likelihood of developing Alzheimer's disease and dying from it; consequently, many of those with *E4* do not generally survive into their nineties. Although *E4* may be one of many potential markers of increased mortality risk, its value as a predictor of Alzheimer's disease has not been proved.

Genes may provide the blueprint for how long a person might live. In effect, they can be considered indicators of how well a person can cope with disease. As such, genes help to determine two interrelated properties that influence aging: adaptive capacity and functional reserve. Adaptive capacity is a person's ability to overcome a disease or injury or to cope with such stresses effectively. Functional reserve refers to how much of an organ is required for its adequate performance. Obviously, one's adaptive capacity depends in part on the body's functional reserve, because the ability to deal with disease requires the proper functioning of organs.

The importance of these two characteristics to the survival of many oldest old can be seen in the varying effects that the buildup of neurofibrillary tangles has on cognition. Neurofibrillary tangles describe the web of dead brain cells that occur naturally with aging but appear in abundance in patients with Alzheimer's disease. The number of tangles that can accumulate before signs of Alzheimer's disease emerge varies among patients. For example, an autopsy revealed that a 103-year-old man who displayed few outward signs of Alzheimer's disease had a level of neurofibrillary tangles that in a younger brain



LISA BURNETT

HIGH THRESHOLD for acquiring chronic diseases and a slower aging process may help explain why the oldest old often survive in good health, according to the author's theoretical model. In most people, tissue damage resulting from disease processes occurs relatively rapidly (*dark blue*). Also, their disease threshold becomes lower quickly with age (*light blue*), so the symptoms of age-related diseases appear by age 80 or so (*a*). Hardy individuals who age slowly (*dark orange*) and have a higher threshold for disease (*light orange*) become symptomatic much later (*b*), if at all.

ANGELINA STRANDEL, age 101, advises those who aspire to her age to "watch your calories and keep away from greasy food." Strandel also indicates that although she has dealt with much turmoil in her life, she does not let the stress get to her. She now lives with her son; longevity seems to run in the family—Strandel's sister lived to age 100.

ERIK LEIGH SIMMONS

VICTORIA ENGER *MSU News Service*

HERBERT KIRK, age 99, graduated in 1993 from Montana State University with a bachelor's degree in art. He is shown here with his senior-thesis sculpture project. Kirk attributes his longevity to exercising: he played tennis into his eighties and still loves to run. When he was 95, Kirk won two gold medals (in 800-meter and five-kilometer races) and one silver medal (in a 200-meter race) at an international seniors' track meet in Helsinki, Finland.

would indicate the patient was probably demented. Presumably, the older man had an excess reserve of brain function that allowed him to compensate for the process that was damaging his brain. Perhaps people who have a slow buildup of tangles and a high tolerance for them can remain mentally intact for a long time, showing overt signs of Alzheimer's disease only very late in life, if at all.

New Thoughts on Aging

The discovery that many people over age 95 are in good shape may mean that future planning for the health care of the oldest old will need to be revised. Much of that planning is based on the theory that although lethal conditions might be postponed as medical technology improves, the incidence of degenerative diseases will be unaltered. The theory predicts that the oldest old will keep suffering from more disease and chronic disability than do people in their eighties. If this theory were true, then the continued increase in the size of the oldest old population would portend a significant burden of poor health among the oldest members of our society.

The emerging data, however, fit better with an opposing theory. James F. Fries of Stanford University has proposed that better ways of life and medical advances will compress morbidity, mortality and disability into a shorter period. Thus, the onset both of major fatal diseases (heart disease, cancer, stroke and Alzheimer's disease) and of age-associated debilitating diseases (degenerative joint disease, sensory impairments and benign memory loss) would be postponed. Consistent with Fries's hypothesis, robust centenarians often have a relatively short period of infirmity before death. Although cause-of-death statistics for centenarians are sparse, available information suggests that the usual causes are acute illnesses such as pneumonia, as opposed to longstanding lethal conditions. In some ways, then, the oldest old resemble Fries's image of the future; perhaps they represent the rare individuals who can already resist disease on their own, without the help of advanced medical science.

Madame Jeanne Calment of Arles, France, will turn 120 in February, making her very likely the oldest living person today whose age has been verified. Some speculate that she may be the oldest living person ever. Most of us do not even come close to this age—those of us with Methuselean aspirations are up against incredible odds. But recent research on the oldest old has prompted new thinking about the biology of aging. Genetic, biochemical and epidemiological studies should reveal exactly why some people possess resistance to debilitating conditions—and may offer ways to increase that ability in a broader swath of the population. Further, the research implies, happily, that as the oldest old become more numerous, they may not become a massive drain on the economy. Counter to prevalent theories of aging, many people in their late nineties or hundreds lead active, healthy lives. If they represent a "survival of the fittest" cohort, the time may have come to abandon our past perceptions of our oldest citizens.

FURTHER READING

AGING, NATURAL DEATH AND THE COMPRESSION OF MORBIDITY. James F. Fries in *New England Journal of Medicine,* Vol. 303, No. 3, pages 130–135; July 17, 1980.

MITOCHONDRIAL GENETICS: A PARADIGM FOR AGING AND DEGENERATIVE DISEASES? Douglas C. Wallace in *Science,* Vol. 256, pages 628–632; May 1, 1992.

THE OLDEST OLD. Richard M. Suzman, David P. Willis and Kenneth G. Manton. Oxford University Press, 1992.

SLOWING OF MORTALITY RATES AT OLDER AGES IN LARGE MEDFLY COHORTS. James R. Carey, Pablo Liedo, Dina Orozco and James W. Vaupel in *Science,* Vol. 258, pages 457–461; October 16, 1992.

ESTIMATES OF CHANGE IN CHRONIC DISABILITY AND INSTITUTIONAL INCIDENCE AND PREVALENCE RATES IN THE U.S. ELDERLY POPULATION FROM THE 1982, 1984 AND 1989 NATIONAL LONG TERM CARE SURVEY. Kenneth G. Manton, Lawrence S. Corder and Eric Stallard in *Journal of Gerontology: Social Sciences,* Vol. 48, No. 4, pages S153–S166; July 1993.

The Birth and Death of Nova V1974 Cygni

The brightest nova in 17 years answered many questions during its life and raised more in death

by Sumner Starrfield and Steven N. Shore

Never has a nova been watched by so many astronomers with so many instruments. Since it exploded in early 1992, nova V1974 Cygni has been recorded in x-rays through radio waves and from the ground, the air, earth orbit and beyond. In its first year the eruption spewed out glowing gases that evolved just as models had predicted 20 years earlier. In its second year the nova, quite unexpectedly, died. We are still trying to understand the short life of V1974 Cygni.

Peter Collins, an amateur astronomer in Boulder, Colo., first spotted the explosion in the early morning of February 19, 1992. Within hours of his report, we looked at the nova with the *International Ultraviolet Explorer* (*IUE*) satellite. We caught it in the "fireball" stage—familiar from photographs of hydrogen bomb explosions, when the gases are first expanding. Before long, it became the only nova to be seen both in birth and in death. In late 1993 the low-energy x-rays coming from the nova's core ceased, indicating that the nuclear explosion had run out of fuel.

SUMNER STARRFIELD and STEVEN N. SHORE have a particular fascination with stellar explosions. Starrfield obtained his Ph.D. at the University of California, Los Angeles, and since 1972 has been a professor at Arizona State University. Starrfield is the principal investigator for observing novae with the *International Ultraviolet Explorer* and the *Compton Gamma Ray Observatory* satellites and runs an electronic-mail network for rapidly disseminating news about novae and supernovae (nova@nova.la.asu.edu). Shore received his Ph.D. in 1978 from the University of Toronto and now chairs the department of physics and astronomy at Indiana University at South Bend. He observed V1974 Cygni with the Goddard High Resolution Spectrograph, which is on board the *Hubble Space Telescope.* Shore serves on the editorial board of the *Encyclopedia of Physical Science and Technology, Comments in Astrophysics* and the *Skeptical Inquirer.*

V1974 Cygni confirmed many of our ideas about novae—such as how the ejected gases evolve—but also destroyed many. It threw off about 10 times more matter than had been anticipated, part of it in the form of dense knots and filaments. The knots may hold the key to the cause of the excess mass. They point to turbulent processes that may have dredged up material from the nova's core. Although we have been forced to rethink many details of how novae evolve, the most essential elements of the original picture remain intact.

In 1892 the nova T Aurigae became the first to be recognized as an explosion, from the peculiarities in its spectrum as compared with normal stars. Since then, scientists have found and studied one or two novae each year. A "naked eye" nova, such as V1974 Cygni, bright enough to be easily visible to the unaided eye, appears perhaps once in a decade.

About 40 years ago a picture of how novae occur began to fall into place. In 1954 Merle F. Walker, then at the Mount Wilson and Palomar Observatories, discovered that the old nova DQ Herculis (which exploded in 1934) is a system of two orbiting stars. One of the stars in the binary system conveniently passes in front of the other, allowing astronomers to measure the time the two stars take to orbit each other. The period turns out to be extraordinarily short—four hours and 39 minutes. One star is also very small; we now know it to be a white dwarf.

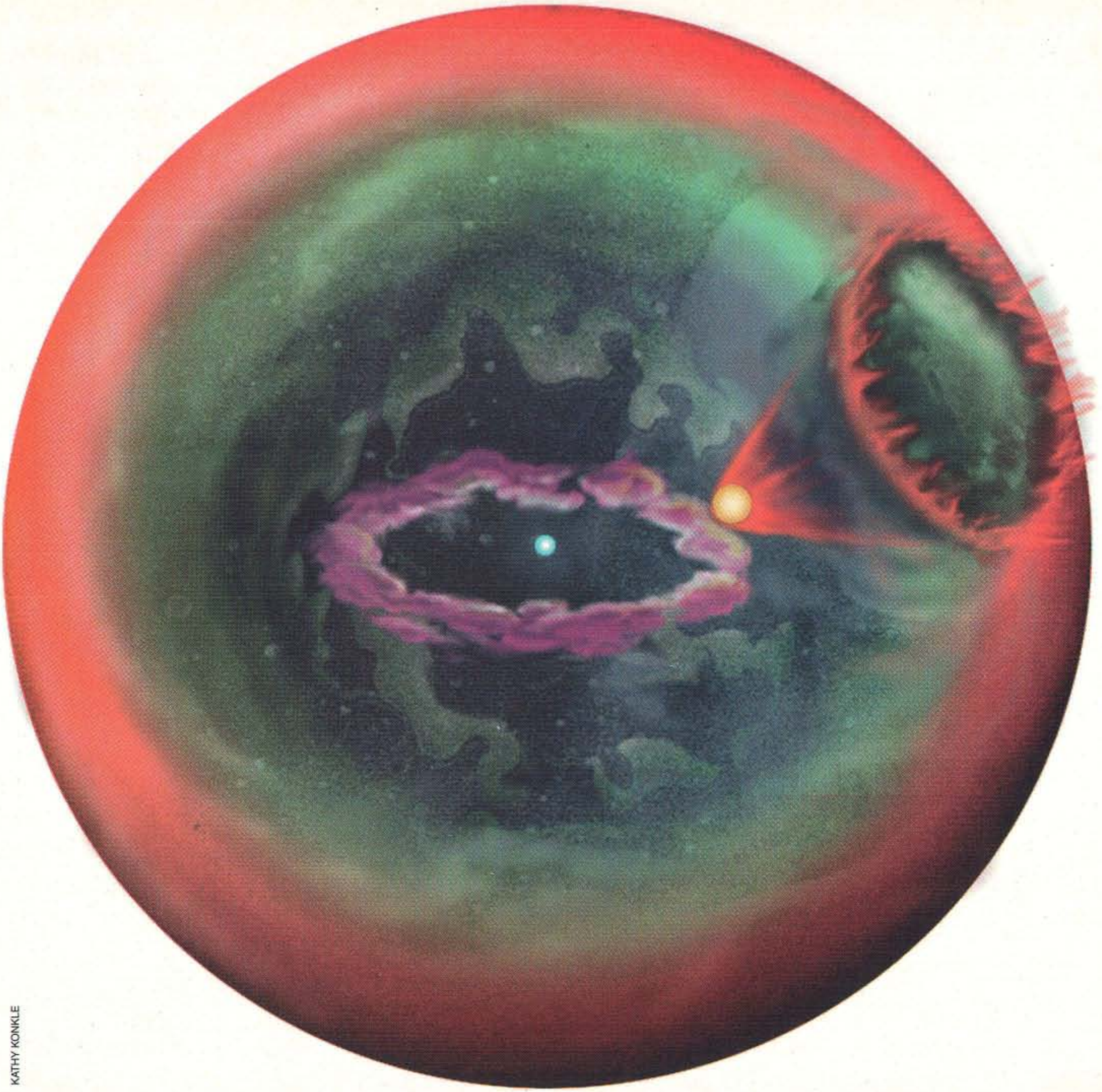
White dwarfs, the end product of stellar evolution, have as much matter as the sun within a volume no larger than the earth's. Robert P. Kraft, also then at the Mount Wilson and Palomar Observatories, showed that other old novae are closely orbiting binary systems. In all these novae, one star was relatively large and unevolved, and the other was a white dwarf. But how can a white dwarf that has no remaining nuclear fuel, along with a stable companion star, give rise to an explosion 10,000 times brighter than the sun? It turns out that each star inexorably alters the other's development.

Calamitous Company

A nova system begins as a widely separated binary, in which one star is more massive than the other. The massive star evolves faster, fusing its hydrogen into helium through the "CNO" cycle of nuclear reactions, which involves carbon, nitrogen and oxygen. At the end of this stage the star becomes a red giant. Its surface swells, engulfing the smaller star. Meanwhile the more massive star fuses the helium in its core to carbon and oxygen.

The stars continue to orbit each other within the common gaseous envelope, losing orbital energy and angular momentum to the gas. As a result, gas is expelled from the system, and the two stars spiral in toward each other. Eventually all the material extending from the massive star past the smaller star is lost. At the end of this "common envelope" evolution, the distantly orbiting stars have become a close binary system. The massive star, having used up all its fuel, has transformed into a compact white dwarf. Its companion has remained relatively unchanged.

Suppose the stars are initially even more widely separated, and the more massive star began its life with about eight to 12 solar masses. Then the latter star can further fuse its core carbon into magnesium and neon. The white



KATHY KONKLE

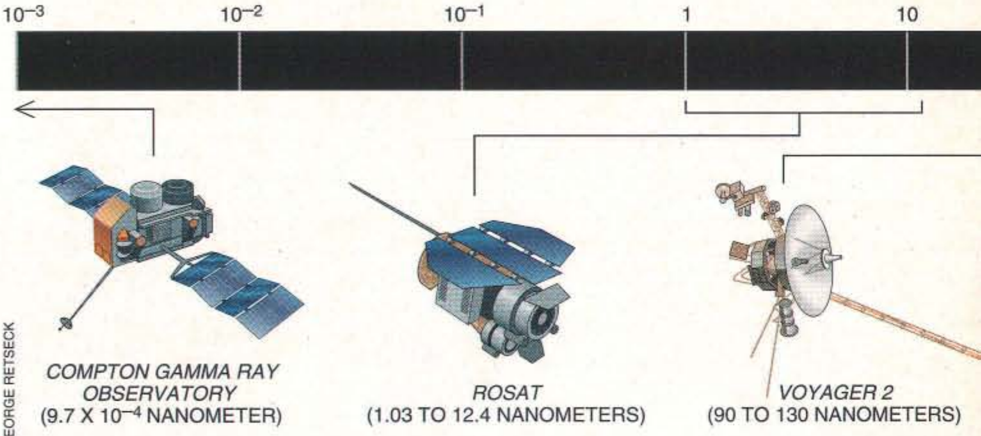
EXPLODING SURFACE of a white dwarf swirls around its companion star in a fireball. Deep within the emitted gases are clumps of dense material (*pink*), probably from the white dwarf's interior. Radiation from residual nuclear reactions on the white dwarf (*center*) continue to heat these clumps until the nova burns out. The ejected gases glow for years.

dwarf it eventually becomes is made of these heavier elements, rather than just carbon and oxygen (a CO nova), and is called an ONeMg white dwarf.

Kraft also made the crucial discovery that the companion star is losing gas. After swirling around in an accretion disk, the hydrogen-rich gas falls onto the surface of the white dwarf. In 1972 one of us (Starrfield, then at the IBM Thomas J. Watson Research Center), along with Warren M. Sparks, then at the National Aeronautics and Space Administration Goddard Space Flight Center, James W. Truran, then at Yeshiva University, and G. Siegfried Kutter, then at the University of Virginia, developed computer simulations that showed how the accreted gas triggers the subsequent explosion.

The intense gravity on the white dwarf compresses the gas as it falls in. If an amount of gas 100 times more massive than the earth accumulates on the white dwarf's surface, then the density in the bottom layer becomes more than 10,000 grams per cubic centimeter. (The density of water is one gram per cubic centimeter.) Because the gas is compressed, its temperature rises to a few million kelvins. The process of accumulation also mixes material from the core of the white dwarf into the overlying and infalling layers, thereby changing their composition.

Under these conditions, the hydrogen nuclei fuse into helium and release energy, by the same CNO nuclear reactions that power normal stars. The ma-


10^{-3}
10^{-2}
10^{-1}
1
10
COMPTON GAMMA RAY OBSERVATORY
(9.7 X 10^{-4} NANOMETER)
ROSAT
(1.03 TO 12.4 NANOMETERS)
VOYAGER 2
(90 TO 130 NANOMETERS)

GEORGE RETSECK

DIVERSE INSTRUMENTS were used to study electromagnetic radiation of different wavelengths emitted by V1974 Cygni. The *Compton Gamma Ray Observatory* searched for photons emitted by the sodium isotope ^{22}Na (and found none). The *ROSAT* satellite detected x-rays coming from the burning core; the cessation of

terial becomes even hotter, so that the fusion proceeds faster, creating runaway thermonuclear reactions like those in a hydrogen bomb.

If the gas were normal, then it would now expand and cool, thereby shutting off the fusion. But the material on a white dwarf behaves in a peculiar manner described by quantum mechanics. It is packed together so tightly that the electrons, which are unable to interpenetrate, become the source of pressure. Unlike an ordinary gas, the material heats up but cannot expand and cool. Nor can radiation carry away the heat fast enough.

The carbon and oxygen mixed in from the core catalyze the CNO cycle and thus speed up the fusion—eventually causing an explosion. The rate of the nuclear reactions depends very sensitively on the temperature, becoming 10^{16} to 10^{18} times faster when the temperature increases by a factor of 10. As the temperature deep within the accreted layers grows to more than 30 million kelvins, the material starts mixing turbulently with the zones above. The mixed region grows toward the surface, carrying with it both heat and nuclei from the interior. Within minutes the surface layers explode into space. They carry along with them fusion products and elements from the dwarf's core, accompanied by a tremendous increase in brightness.

Burning Out

The first few minutes of a nova explosion have never been observed. Our simulations predict that the surface temperature can exceed one million kelvins and that the hot gases are blown away at more than 5,000 kilometers per second. Because its volume increases suddenly, the gas cools. In a few hours the radiation it emits shifts from being primarily in x-rays to the lower-energy ultraviolet. At the same time, the surface area of the gas increases, making the nova brighter even as it becomes cooler. A spectacular transformation ensues.

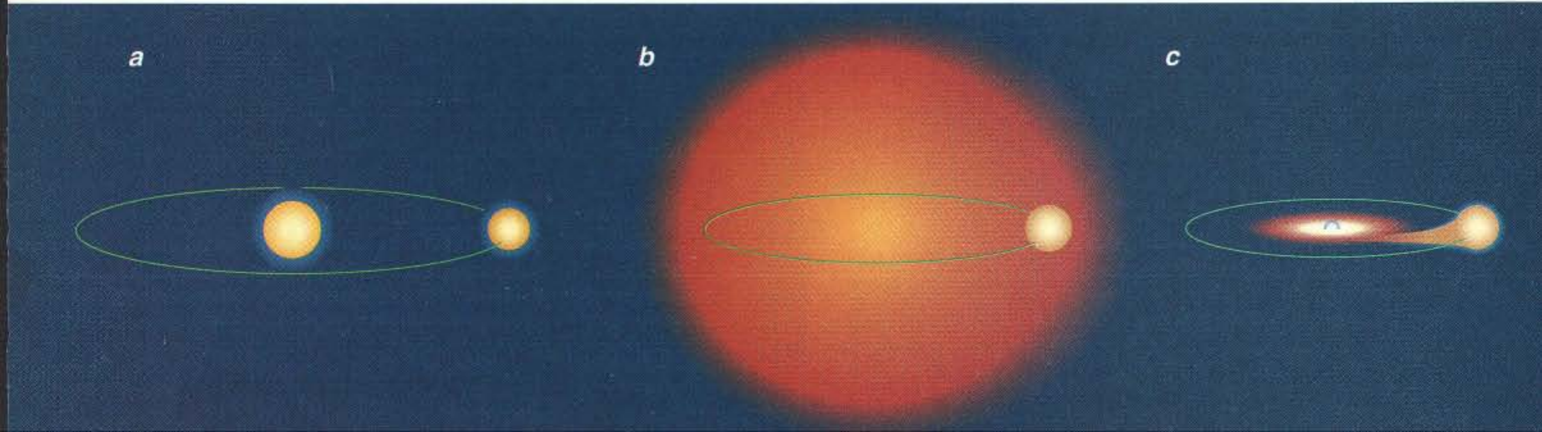
Initially, the expanding shell consists of a hot, dense gas of electrons and ions—atoms missing one or more electrons. This gas is reasonably transparent. But as it expands, its temperature drops below 10,000 kelvins. The electrons start to recombine with the ions to form atoms that are missing either one or no electrons. These atoms have many energy levels and can absorb tens of millions of individual wavelengths of light.

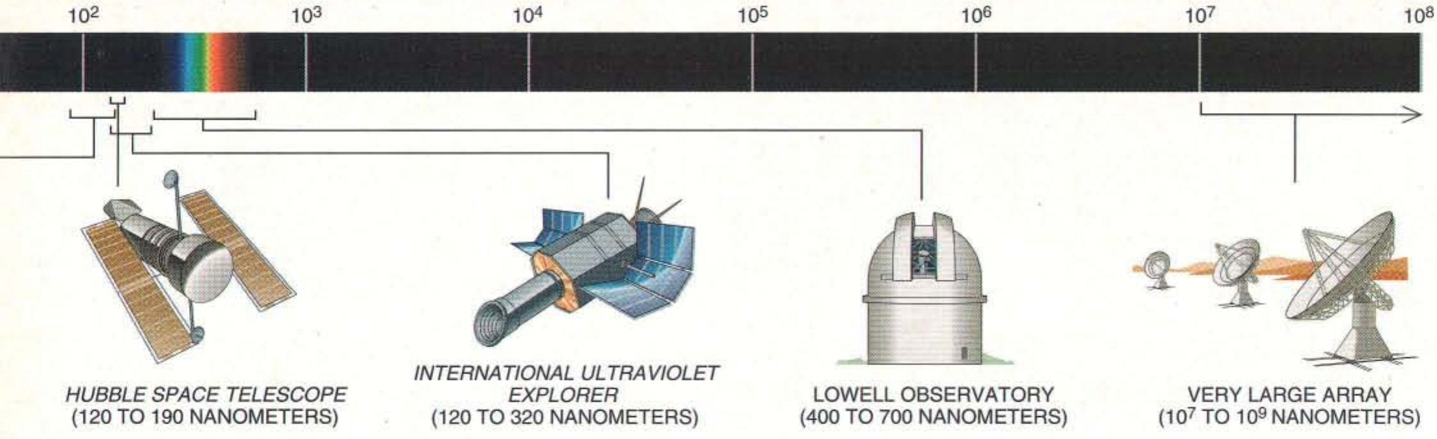
The most important absorbers have atomic numbers around 26, that of iron. The spectrum of light that they can absorb is extremely complex. These ions and atoms block most of the energy being radiated in the ultraviolet, which is where most of the energy is emitted at this phase. When we first studied this phase, with Peter H. Hauschildt of Arizona State University and other collaborators, we called it the iron curtain. The energy absorbed by the curtain is reemitted at longer—optical and infrared—wavelengths.

The iron curtain was vividly confirmed by our first observations of V1974 Cygni. Within hours of its discovery George Sonneborn of the Goddard Space Flight Center activated our Target of Opportunity program, which allows us to observe immediately with the *IUE* satellite when a bright nova occurs. Pointing the wonderfully maneuverable satellite at the nova, he obtained a series of ultraviolet spectra.

Within an hour we were able to ob-

NOVA SYSTEM begins as a pair of widely separated orbiting stars (*a*). The more massive star evolves faster, becoming a red giant and enveloping the smaller star (*b*). The stars lose angular momentum to the gas and spiral in toward each other while the gas is expelled. Eventually they form a close binary system in which the remaining core of the red giant, having used up all its fuel, has become a white dwarf. The less massive star now sheds matter, which first forms an accretion disk (*c*). Falling onto the white dwarf's surface, this material is compressed by the high gravitational field. Then a runaway thermonuclear reaction—a nova explosion—takes place (*d*), stripping most of the accreted material off the white dwarf (*e*). It can, however, accrete fuel from its neighbor again, cycling through the steps from *c* to *e* many times.





these rays signaled the nova's demise. *Voyager 2,* then beyond Neptune, observed far ultraviolet radiation, the first for a nova. The *International Ultraviolet Explorer* captured the explosion in its early fireball stage. The *Hubble Space Telescope* revealed clumps within the ejected gases. The 72-inch telescope at Lowell Observatory in Flagstaff, N.M., recorded optical light, and the Very Large Array in Socorro, N.M., detected radio emissions that confirmed the presence of clumps.

serve that the nova's ultraviolet brightness had dropped slightly [*see bottom illustration on next page*] and that its optical brightness had risen. Astronomical change is measured, as a rule, in billions of years; it is rare to see evolution on such a short timescale. During the next two days, the ultraviolet radiation dropped to 3 percent of its original value. All the while the nova became optically brighter. As soon as the visual brightness peaked, the ultraviolet emissions bottomed out and began to climb.

The recovery comes from a second change in ionization. As the gas expands, its density drops. Then the iron group elements once again become ionized and hence transparent. Radiation now flows from the interior, enhancing the ionization and in turn the transparency. In effect, the iron curtain lifts, and ultraviolet light from the hot, deep layers penetrates through the outer layers. Within two months the ultraviolet brightness had climbed back up to its original value.

At the same time as the ultraviolet brightness increased, the visual brightness of the nova declined. The total (bolometric) brightness of the underlying star remained, however, virtually unchanged. This "constant bolometric luminosity" phase, predicted by our 1972 simulations, was finally confirmed in detail by observations of V1974 Cygni.

Anticipating that the radiation peak would continue to move toward shorter wavelengths, Ronald S. Polidan of the Goddard Space Flight Center requested that *Voyager 2,* then flying beyond the orbit of Neptune, observe the spectra of V1974 Cygni. On April 27, 1992, the spacecraft detected the nova—the first to be seen in the far ultraviolet. Its brightness in this wavelength range increased during the observations.

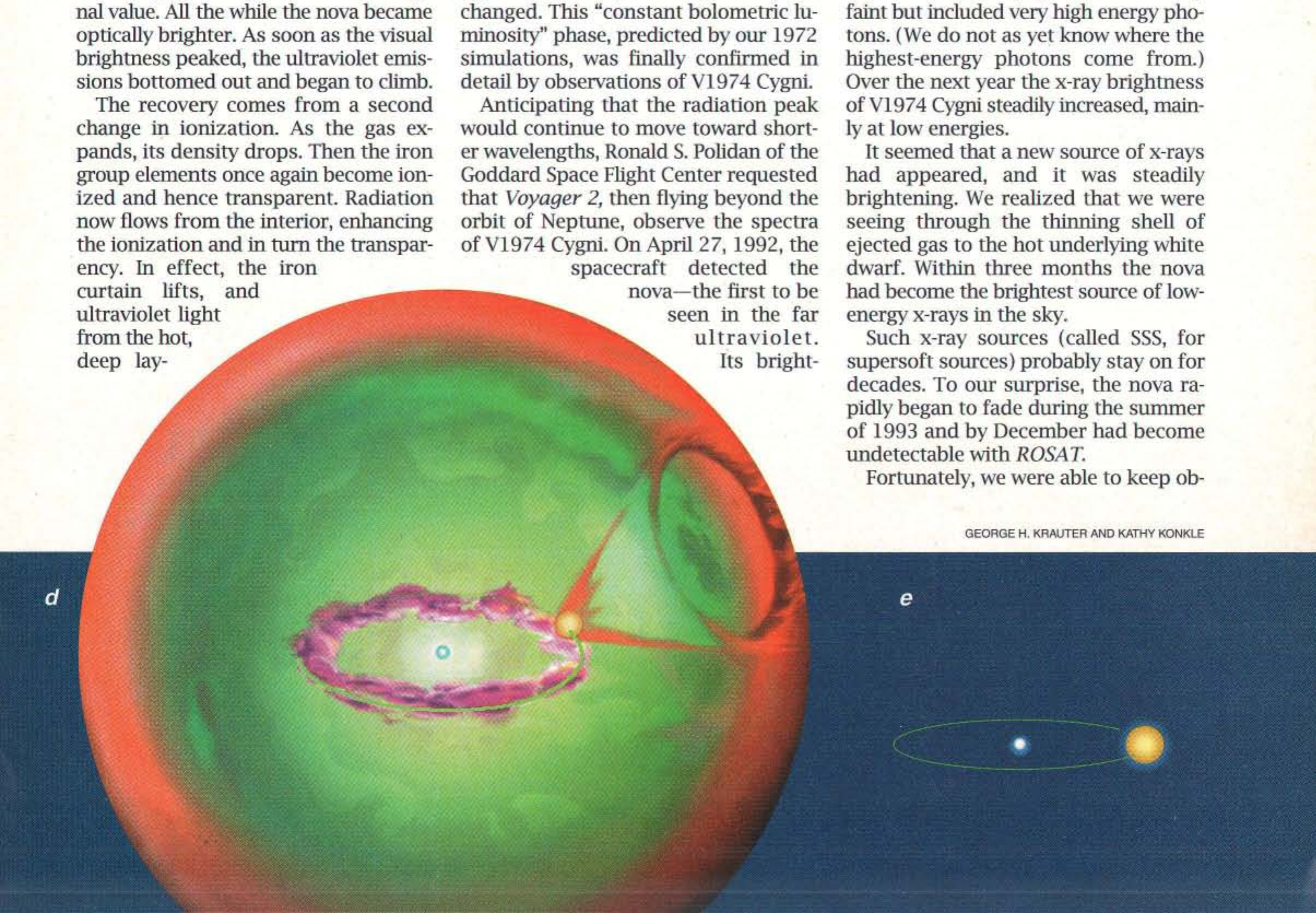
The radiation peak continued to shift into shorter wavelengths. Using the *ROSAT* satellite, Joachim Krautter of Heidelberg Observatory, Hakki Ögelman of the University of Wisconsin and Starrfield started observing the nova on April 22, 1992. The x-ray spectrum was very faint but included very high energy photons. (We do not as yet know where the highest-energy photons come from.) Over the next year the x-ray brightness of V1974 Cygni steadily increased, mainly at low energies.

It seemed that a new source of x-rays had appeared, and it was steadily brightening. We realized that we were seeing through the thinning shell of ejected gas to the hot underlying white dwarf. Within three months the nova had become the brightest source of low-energy x-rays in the sky.

Such x-ray sources (called SSS, for supersoft sources) probably stay on for decades. To our surprise, the nova rapidly began to fade during the summer of 1993 and by December had become undetectable with *ROSAT.*

Fortunately, we were able to keep ob-

GEORGE H. KRAUTER AND KATHY KONKLE

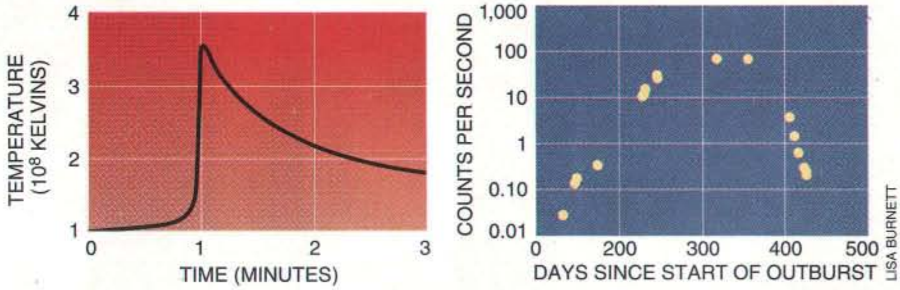


serving with the *IUE.* We found that the amount of highly ionized nitrogen was declining, which meant that the ions were recombining with electrons to form less ionized atoms. Furthermore, nitrogen ions that were missing four electrons were recombining faster than were the ions missing three electrons. Apparently the intense radiation that had been stripping the nitrogen of its electrons had vanished: the x-rays were indeed gone. To us, this absence could mean only that the white dwarf had consumed all its fuel and that the nuclear fusion on its surface had ceased.

The nova outburst had lasted about 18 months. The life span of a nova depends on the mass of the white dwarf that hosts it. A massive white dwarf compresses the accumulated gases more intensely. In that case, fusion starts early, and the fuel runs out quickly, causing the nova's life to be brief. Also, the explosion ejects much less matter than does one on a low-mass white dwarf. According to our models, the short life of V1974 Cygni implies that its mass was 20 to 30 percent greater than that of the sun. The mass of material it expelled should have been about 10^{-5} solar mass. But the amount shown by observations is about 10 times more. We do not understand this discrepancy.

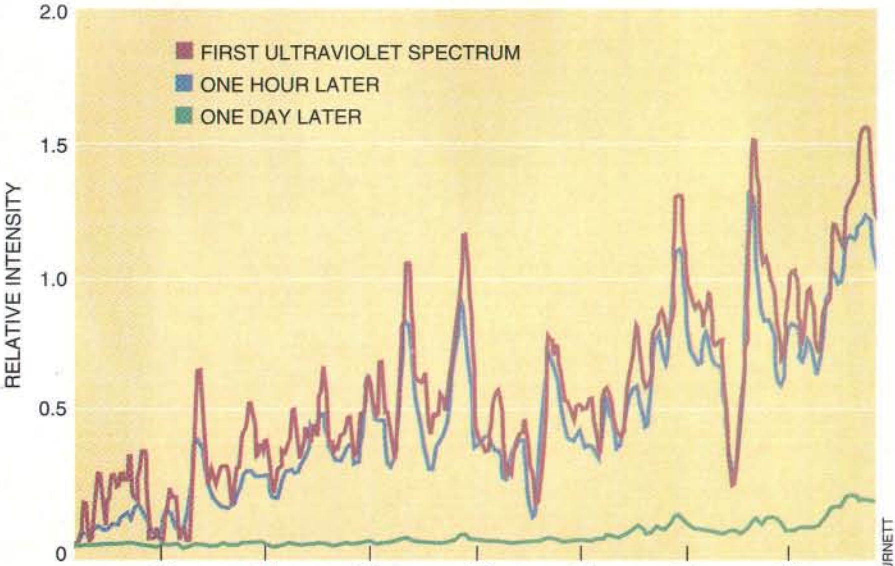
Clumpy Clues

Some hints to the problem may possibly be found in the knots. Our first clear view of the knots was on September 7, 1992, when we observed the nova with the Goddard High Resolution Spectrograph (GHRS) on the *Hubble Space Telescope.* With this powerful instrument we obtained the highest-quality ultraviolet spectra ever for a nova. Each emission line showed evidence that the gas had been expelled in two stages. There was high-velocity gas that had been ejected uniformly and denser, slower-moving clumps.



LISA BURNETT

NUCLEAR REACTIONS make the temperature (*left*) on the surface of a white dwarf rise to almost 400 million kelvins within a minute. As the gases blow out in a nova explosion, their temperature plummets. The x-ray emissions (*right*) observed by *ROSAT* came from the remaining core of the white dwarf. The emissions died off faster than expected, indicating that the white dwarf had used up all its fuel.



LISA BURNETT

IRON CURTAIN falls when expanding gases from a nova cool to form atoms and ions of atomic numbers near that of iron (26). These ions absorb all the ultraviolet light emitted by the nova. The first spectrum (*purple*) taken by the *International Ultraviolet Explorer* shows strong emissions. Within an hour (*blue*) the ultraviolet radiation had fallen, and by the next day (*green*) it was almost gone.

Armed with the high-quality GHRS spectra, we reexamined the earlier data from the *IUE.* The spectra we had gotten just after the iron curtain cleared also displayed the knots. This evidence indicates that the structures had been formed during the explosion. Looking again with the GHRS on April 1, 1993, we found the same clumps we had identified earlier, moving at the same speeds. The faster material had largely vanished, so we were now seeing completely through the ejected gas.

The spatial structure of the shell was resolved in December 1992 with the Very Large Array radio telescope by Robert M. Hjellming of the National Radio Astronomy Observatory. His radio images of the expanding material confirmed our analysis of the knots. According to images taken with the *Hubble Space Telescope,* the shell was resolved in the ultraviolet by May 1993.

This is the first time we have ever had such a clear view so early of the debris from a nova explosion. The knots appear to be deeply embedded within the ejecta. Now we need to understand what caused them and what they are made of.

The first indications of the composition of the ejecta came around April 1, 1992, when the iron curtain had finally lifted, leaving an intense spectrum with bright emission lines from carbon, nitrogen, oxygen and other abundant elements. Previously, we had encountered emission lines of this kind only in novae that took place on massive ONeMg white dwarfs. We conjectured that V1974 Cygni, too, belonged to this class. The idea also occurred to Thomas L. Hayward of Cornell University and Robert D. Gehrz of the University of Minnesota and their collaborators, who had just obtained infrared spectra of the nova using the five-meter telescope on Mount Palomar. They found the characteristic 12-micron line emitted by ionized neon. This line is normally very weak or absent in CO novae but is strong in ONeMg novae.

In the fall of 1993 the gases thinned enough so that Scott Austin of Arizona State University, R. Mark Wagner of Ohio State University and the two of us could finally use the optical and ultraviolet spectra to determine the chemical abundances of the debris. (While the

HAROLD EDGERTON

FIREBALL billows out a fraction of a second after an atomic bomb explosion at a Nevada test site. Its structure is very similar to the fireball from a nova. This photograph from the 1950s was taken by automatic instruments situated 20 miles away. In the foreground are Joshua trees, about to be incinerated. The intense heat melted desert sand into glass.

gas was dense, the atoms collided with one another, thus complicating the spectra greatly.) We found large quantities of elements from the core. In a paper we are preparing for the *Astrophysical Journal,* we report that the expelled material has more than 30 times more oxygen, neon, nitrogen and aluminum than solar material. This composition is similar to that of other ONeMg novae. It suggests that the shell and the core are churned together by turbulent processes not included in our current models, perhaps creating the knots and eventually expelling large fragments of the core.

Riddles Remain

Another, related mystery pertains to the elements synthesized during the explosion. Achim Weiss of the Max Planck Institute for Astrophysics in Garching, Irit Idan and Giora Shaviv of the Technion University in Israel, Truran and Starrfield have calculated that ^{22}Na, an isotope of sodium with mass number 22, should be produced in an ONeMg nova. This isotope is radioactive, with a distinct pattern of gamma-ray emissions.

Our calculations indicate that V1974 Cygni produced large amounts of ^{22}Na. With the *Compton Gamma Ray Observatory,* we searched for the appropriate gamma rays in September 1993—but found none. (To be sure, our observing time was drastically cut by the Target of Opportunity program being invoked for other celestial events.)

All these anomalies tell us that although we have come a long way in understanding nova explosions, we still have much to learn. We understand the thermonuclear reactions that produced the explosion. What is not so clear is the dynamics. Do the shell and the core mix while material is being accreted or during the last stages of the explosion?

Another mystery is the long-term effect of repeated nova outbursts on the evolution of the white dwarf. All nova binary systems go through the cycle of accretion and explosion many times. If parts of the core are shed during each outburst, then the mass of the white dwarf must be decreasing with repeated explosions. Does its mass become ultimately very small, or does something happen to stop any further outbursts?

Because of the brightness and slow evolution of its debris, we will be observing nova V1974 Cygni well into the 21st century. We hope the nova will supply some answers to the questions it has raised.

FURTHER READING

THE CLASSICAL NOVA OUTBURST. Sumner Starrfield in *Multiwavelength Astrophysics.* Edited by France A. Córdova. Cambridge University Press, 1988.

THERMONUCLEAR PROCESSES AND THE CLASSICAL NOVA OUTBURST. Sumner Starrfield in *Classical Novae.* Edited by Mike Bode and Nye Evans. Wiley, 1989.

OBSERVATIONS AND PHYSICAL PROCESSES IN BINARY STARS. Steven N. Shore in *Interacting Binaries.* Edited by S. N. Shore, M. Livio and E.P.J. van den Heuvel. Springer-Verlag, 1994.

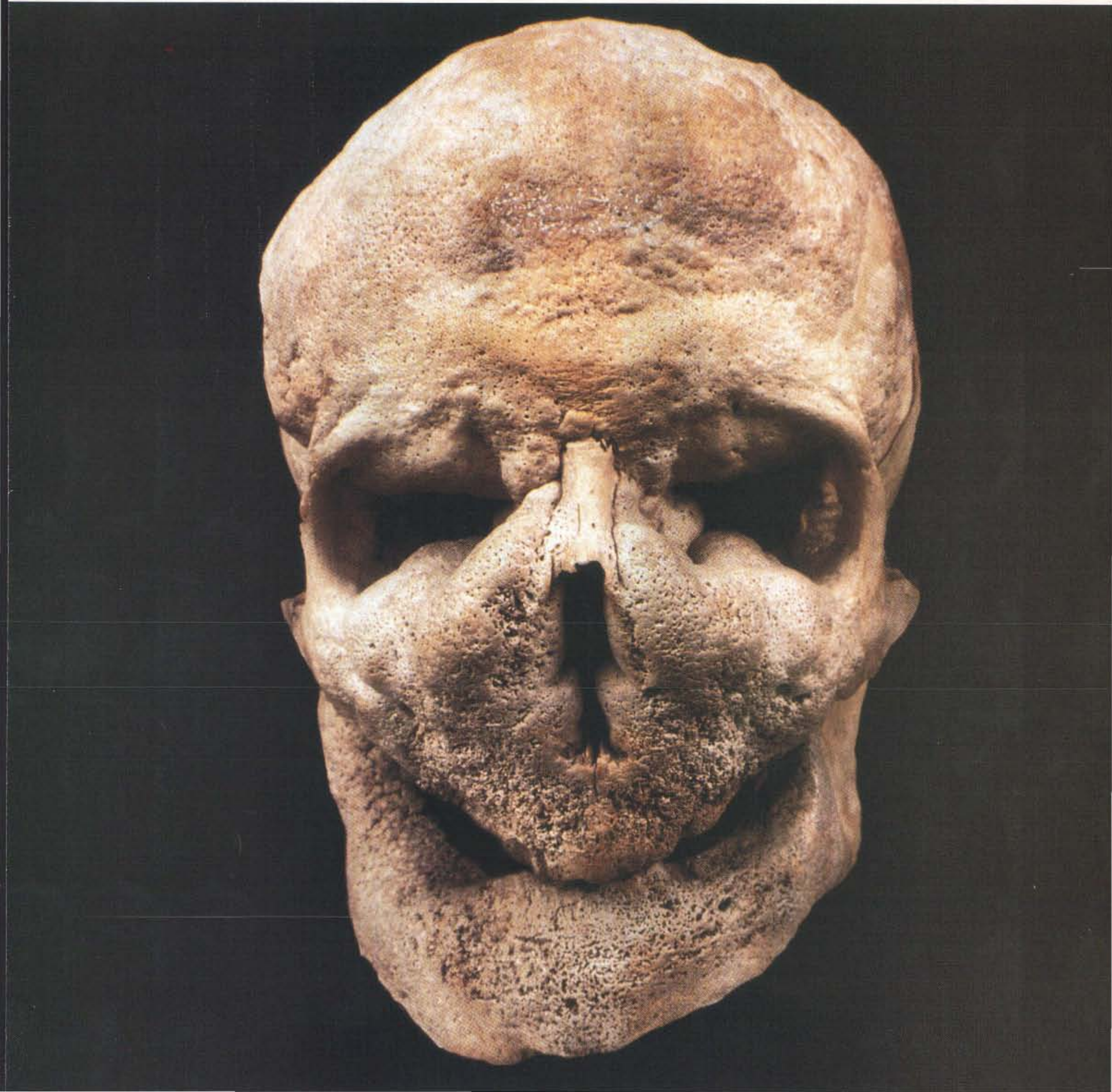
TOPICS IN CLOSE BINARY EVOLUTION. Edward P. J. van den Heuvel in *Interacting Binaries.*

TOPICS IN THE THEORY OF CATACLYSMIC VARIABLES AND X-RAY BINARIES. Mario Livio in *Interacting Binaries.*

Egil's Bones

An Icelandic saga tells of a Viking who had unusual, menacing features, including a skull that could resist blows from an ax. He probably suffered from an ailment called Paget's disease

by Jesse L. Byock

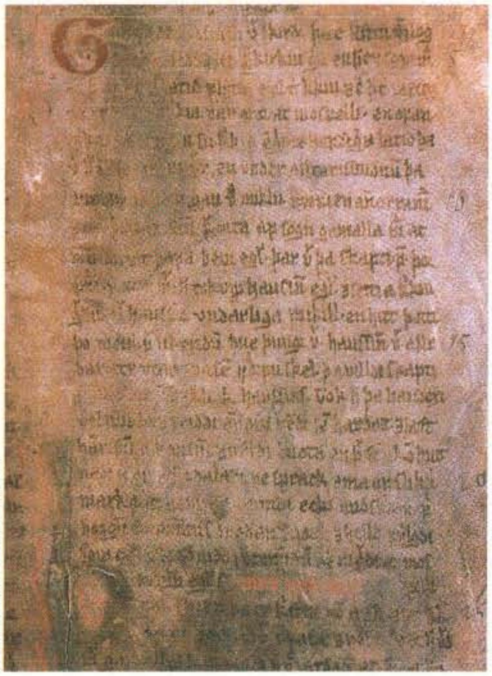


Egil, the son of Skalla-Grim, is the most memorable Viking to appear in the Old Norse sagas. Born in Iceland in the early 10th century, he participated in Viking raids and adventures throughout Norway, Sweden, Denmark, the east Baltic lands, England, Saxony and northern Germany. Fierce, self-willed and violent, Egil Skalla-Grimsson was also a fine poet and a man with a sense of ethics. He epitomizes the Viking urge to travel into the unknown world seeking action and fortune. From Athelstan, king of the Anglo-Saxons, he receives valuable gifts and pledges of friendship, but from Erik Blood-Axe, the Viking ruler of Norway, he hears death threats. Combining courage and brawn with high intelligence, Egil survives war and treachery to live to an old age of 80. He dies among his kinsmen in Iceland in about 990, apparently from natural causes stemming from longevity.

For all Egil's heroic stature, however, there is something deeply troubling about his character. Despite his prowess and secure social status, his temperament, as well as his physical appearance, causes alarm. He is portrayed as an ugly, irritable, brooding individual. In this respect, Egil resembles his father and his grandfather, men described as physically menacing. The saga clearly distinguishes them as physiologically different from their kinsmen, who are depicted as fair and handsome.

What set Egil apart was more than simply a small, personal peculiarity. Through prose and verse, the saga tells us that Egil became deaf, often lost his balance, went blind, suffered from chronically cold feet, endured headaches and experienced bouts of lethargy. Furthermore, the saga describes unusual disfigurements of his skull and facial features. These symptoms suggest that Egil may have suffered from a syndrome that results from a quickening of normal bone replacement. The disease, first diagnosed by Sir James Paget in 1877, runs in families and is uncannily similar to Egil's affliction.

Is it really important to determine whether Egil suffered from Paget's disease? I pondered this question at the beginning of my research and considered it again when I realized that the enigma of Egil lies at a nexus of medical science, history, archaeology and literary analysis. The answer is yes: such a determination does matter. An understanding of Egil's affliction is a critical step in assembling the evidence needed to evaluate the historical accuracy of the Icelandic sagas. Do sagas provide accurate information about a Viking period 250 years before they were written? Or are they merely flights of fancy and fabrications by 13th-century authors? Historians, literary scholars, archaeologists and linguists have all had their say, but science has scarcely played a role in the debate. At times the subject has stirred so much passion that one scholar promised to maintain his view until forced by death to lay down his pen. The argument would change drastically if a new source of information could be found.


STOFNUN ARNA MAGNUSSONAR

SAGA PAGE from the Icelandic manuscript *The Book of Möðruvellir* contains the passage describing Skapti's unearthing of Egil's bones. In the 12th line, one can easily make out: *Þar var þá, Skapti prestr,* meaning "Skapti the priest was there at the time."

For me, that new source lies unexpectedly in the field of modern medicine. Rather than attributing conflicting aspects of Egil's personality to artistic hyperbole, I believe the descriptions stem from the progress of Paget's disease. In breaking tradition to arrive at these conclusions, I frequently have recourse to another science—philology, the historical and comparative study of language and its relation to culture.

Family Stories

The Icelandic sagas constitute one of the largest collections of extant vernacular narratives from medieval times. In 31 major sagas and scores of shorter narratives, these texts recount the travels of the first generations of Norse settlers in Iceland, the major Viking outpost in the North Atlantic. Written in prose and studded with verse, the family sagas are set in the period from 870 to 1030. Unlike myths and fantastic tales, which the Icelanders also produced, the sagas are sober in style. With an often stark realism they detail everyday agrarian and political life and describe adventurous Viking voyages, including those to Greenland and Finland. The crucial question that has remained is whether the sagas are the product of a long oral tradition or the invention of authors after Iceland became literate in the 13th century.

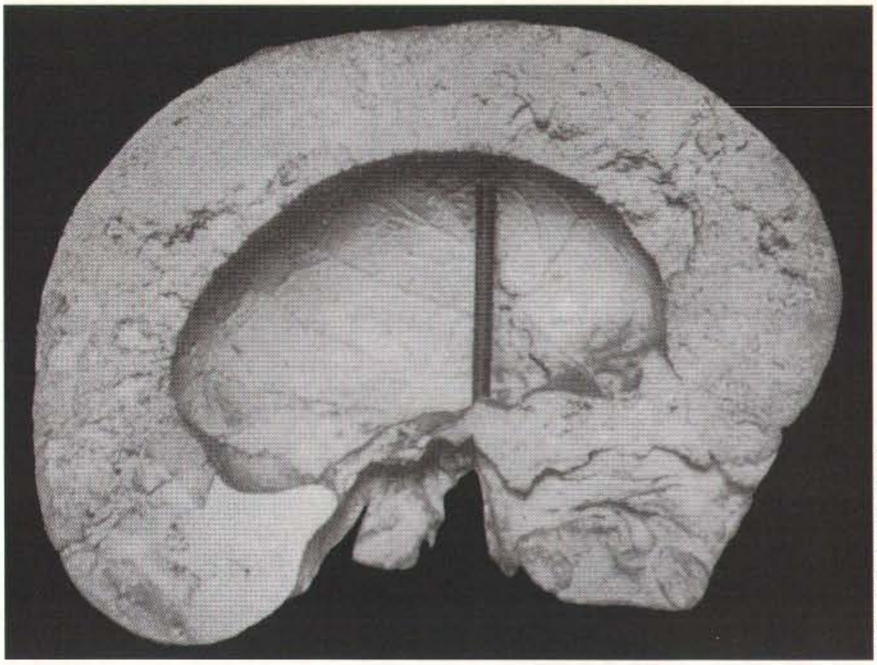
According to *Egil's Saga,* Egil spent his final years with his adopted daughter, Thordis, at the farm of Mosfell, in southwestern Iceland, not far from the present-day capital of Reykjavík. Initially he was buried there in a pagan grave mound. But 10 years later, when Iceland converted to Christianity in 1000, Thordis and her husband, Grim, moved Egil's remains to be interred at a small church built on their farm. About 150 years later a second church was built about 500 meters from the first. Skapti, one of Egil's prominent descendants, exhumed Egil's bones to move them to the new churchyard. The final pages of *Egil's Saga* relate a curious tale about Skapti's findings:

> Under the altar some human bones were found, much bigger than ordinary human bones.... Skapti Thorarinsson, a priest and man of great

JESSE L. BYOCK is professor of Old Norse and medieval Scandinavian studies at the University of California, Los Angeles. He has published three books on Iceland and the sagas—*Medieval Iceland: Society, Sagas, and Power; Feud in the Icelandic Saga;* and *The Saga of the Volsungs: The Norse Epic of Sigurd the Dragon Slayer.* After studying in Sweden, Iceland and France, he received his Ph.D. from Harvard University. A specialist in North Atlantic and Viking studies, he once herded sheep in the northern fjords of Iceland.

ROYAL COLLEGE OF SURGEONS, ENGLAND

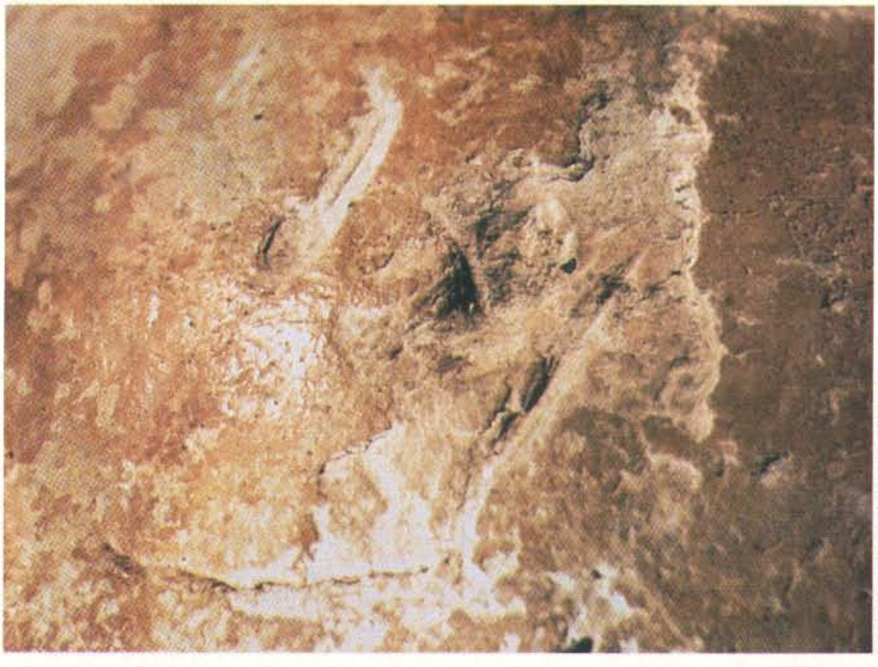
FACIAL BONE that is grossly thickened and symmetrically deformed may result from Paget's disease, a condition in which bone grows drastically out of control. The skull of the warrior Egil, a hero of the Icelandic sagas, may have looked similar to this one of a man who suffered from the disease several centuries ago.



BRITISH JOURNAL OF SURGERY

THICKENED PAGETIC SKULL, shown in cross section, does not intrude into the cranial cavity, so that a patient's intellect remains unimpaired.

> intelligence, was there at the time. He picked up Egil's skull and placed it on the fence of the churchyard. The skull was exceptionally large, and its weight was even more remarkable. It was ridged all over on the outside like a scallop shell. Skapti wanted to find out just how thick the skull was, so he picked up a heavy ax, swung it in one hand and struck as hard as he was able with the hammer side of the ax, trying to break the skull. But the skull neither broke nor dented when hit; it simply turned white at the point of impact, and from that result anybody could guess that it could not have been easily cracked by small fry while still covered by skin and flesh.



ROYAL COLLEGE OF SURGEONS, ENGLAND

WHITENING OF PAGETIC BONE, caused by impact, is apparent in this centuries-old, fossilized parietal bone unearthed in the mid-1700s. A similar whitening occurred when Egil's skull was struck with the hammer side of an ax.

This passage has often been used to exemplify the untrustworthiness of the sagas. No matter how realistic the description may seem to be, everyone, so the argument goes, knows that a 150-year-old skull could not possibly have withstood the blow of Skapti's ax.

Modern medical knowledge, however, suggests that we cannot treat this episode as a purely literary device intended to magnify heroic Viking qualities. *Egil's Saga* precisely describes the skull as "ridged all over on the outside like a scallop shell." The precision is striking because the passage marks the only instance in all Old Norse literature that the otherwise well-known words *hörpuskel* (scallop shell) and *báróttr* (ridged, undulated, waved, corrugated, wrinkled) are used to describe human characteristics. A "scalloped" bone surface, unique in descriptions of Viking heroes, closely matches medical portrayals of Paget's disease. Diagnoses repeatedly list irregularities of the outer skull surface, describing its appearance as corrugated and wavy. Such a feature appears in about one in 15 symptomatic cases.

Physicians have also noted the exceptionally resilient, ivorylike hardness of the afflicted bones. Even the whitening of Egil's skull where Skapti's ax struck is a clear indication of Paget's disease. When subjected to a blow, the soft, pumicelike outer material of the enlarged Pagetic skull gives way to a white, hardened, highly resilient core.

A Helm's-Rock of a Head

In the saga, Egil himself refers to his head in strange ways. In one poem, written in response to a pardon granted him by his sworn enemy, King Erik Blood-Axe, Egil composed this verse:

> *I am not opposed,*
> *Though I may be ugly,*
> *To accepting my helm's-rock*
> *of a head,*
> *From a ruler.*

Paget's disease may have been responsible for Egil's memorable facial appearance, a subject that over the years has engendered a great deal of creative literary interpretation. The saga offers a detailed description of Egil as he sits at a feast in England after battle. He faces King Athelstan, the ruler who he believes owes him compensation for the death of his brother, Thorolf.

> Egil sat upright, but his head hung forward. Egil was marked by prominent features. He had a broad forehead and large eyebrows, a nose that was not long but enormously thick, and lips that, seen through his beard, were both wide and long. He had a

remarkably broad chin, and this largeness continued throughout the jawbone. He was thick-necked and broad-shouldered and, more so than other men, hard-looking and fierce when angry. Well built and taller than others, he had thick wolf-gray hair but was early bald. While he sat as was written above, he jerked one eyebrow down to his chin and lifted the other one up into his hairline; Egil was black-eyed with eyebrows joined. He refused to drink, although drink was brought to him, but alternately he jerked his eyebrows up and down.

King Athelstan does not overlook Egil's threatening stance. Acting to defuse the Viking's anger, Athelstan offers Egil a liberal payment in compensation for his brother's death and thus wins his loyalty.

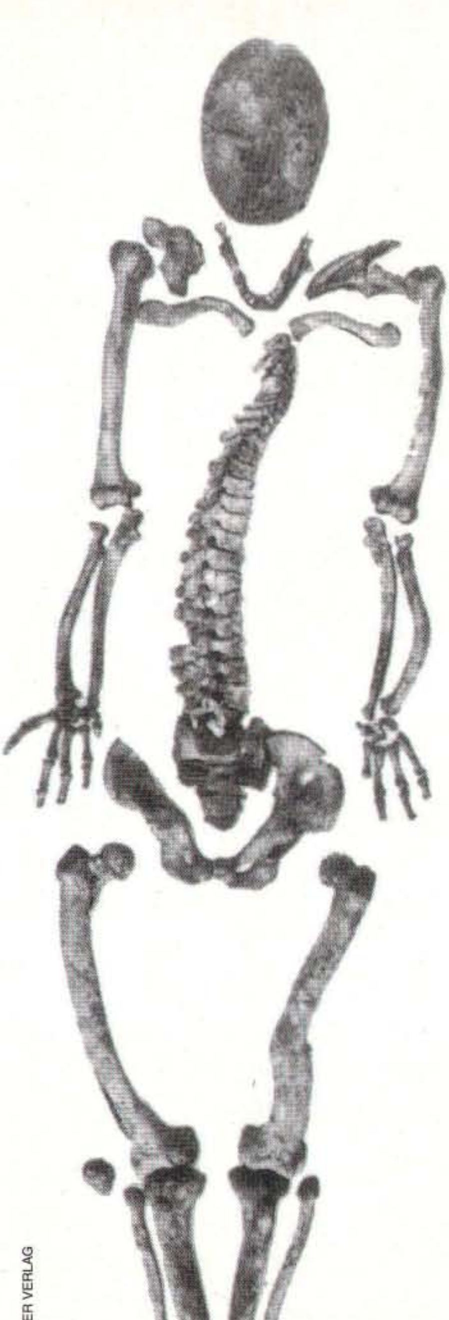
Although literary license may be assumed, it is curious and highly unusual for the physical features of a saga protagonist to be portrayed in so grotesque a manner, unless the writer was reporting a well-known story. Distortion and hardening of the cranium, changes that are characteristic of Paget's, may lead to leontiasis ossea, or cranial hyperostosis. In this condition, the facial bones thicken, giving an individual a lionlike appearance. This pathology, which may occur as early as the first two decades of life, closely fits the descriptions of Egil. As for the bizarre mobility of the eyebrows, it is conceivable that a person as menacing as Egil learned to exploit his facial distortion and was remembered for its outrageous effect.

Corroborating evidence for a Pagetic diagnosis comes from the saga's description of the problems Egil had in his old age. These difficulties—including loss of balance, hearing and sight, chill in the extremities, headaches and the phenomenon described as a hanging, swaying head—are all major symptoms of advanced Paget's disease. According to the saga, the elderly Egil, after moving in with his son-in-law, Grim, at Mosfell, was walking outside one day when he stumbled and fell. Some of the women who saw this laughed:

"You're really finished, now, Egil," they said, "when you fall without being pushed."

"The women didn't laugh so much when we were younger," said Grim. Egil made this verse:

The horse of the necklace sways,
My bald head bangs when I fall;
My piece's soft and clammy
And I can't hear when they call.



HANS HUBER VERLAG

TENTH-CENTURY SKELETON of an Anglo-Saxon male shows the ravages of extensive Paget's disease, such as the curved spine and the thickened and deformed long bones of the limbs.

Why would people remember this poem about a head that "sways" and other physical difficulties? One reason is that the utterance is a powerful example of Old Norse verse, encasing personal emotion in a complex and colorful word puzzle. In Viking times, verse was viewed as the gift of Odin, and poetic skill was highly respected. Here the lines reflect the aged warrior's still agile ability to turn physical disorders into memorable imagery.

Old Norse poetry was a game of puzzles, which, once the rules are understood, supplies us with critical information. The first line carries the understanding: "I have swayings of the neck." In building this image, the author fashioned a Norse poetic circumlocution called a kenning. Kennings are stylistically similar to certain English metaphors, such as calling a camel a ship of the desert. The saga's kenning, *helsis valr* (the horse of the necklace), means the neck. The word "swayings" is built on the verb *váfa,* "to sway or dangle while hanging." Thus, the line in the verse refers to a neck bent under the weight of a head that wobbles.

A drooping, swaying head is not a standard feature of old age; so, too, the graphic description "swayings of the neck" is by no means a common usage in Old Icelandic poetry. I conducted a computer search and found no other occurrences of this combination, so the poet is clearly describing a condition that is unusual and highly personal.

The saga further narrates that Egil becomes blind and is humiliated in his old age by his lethargy and his craving for warmth—all symptoms of Paget's:

Egil became totally blind. One day in winter when the weather was cold, he went up to the fire to warm himself.... "On your feet!" said the woman. "Get back to your place and let us do our work." Egil stood up, walked over to his seat, and made this verse:

I flounder blind by the fireside,
Ask females for mercy,
Bitter the battle
On my brow-plains.

In this verse the Icelandic words for "brow-plains" (*hvarma hnitvellir*) mean the part of the face where the eyes meet or are located. The passage is thus unclear as to whether the words refer to the eyes themselves or to the area of the eyes, including the part behind and around the eye sockets. If the former, the words again mean Egil's blindness. If the latter, the phrase expresses the notion that "I bear pain where the eyes meet," suggesting that Egil has headaches. Possibly, both interpretations were intended.

Egil's headaches and chills are consistent with his other symptoms. Victims of Paget's disease sometimes have headaches caused by the pressure of enlarged vertebrae on the spinal cord. They also show a high incidence of arteriosclerosis and heart damage. Attendant circulatory problems, particularly coldness in hands and feet, develop as the heart is overtaxed and blood is diverted from the extremities in order to support the rapid bone remodeling.

Cold Feet, Cold Women

Another of Egil's laments supplies further information about his chills and cold feet—and of his ability to create clever wordplays.

Two feet I have,
Cold widows.
These frigid crones
Need a flame.

In Icelandic, the words are

Eigum ekkjur
allkaldar tvær,
en þær konur
þurfa blossa.

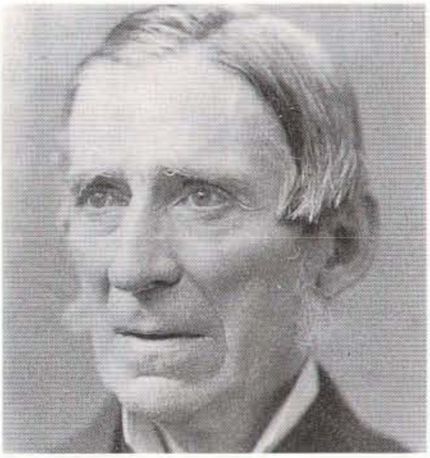
Here the poet is skillfully playing on an understood double entendre. In unraveling the puzzle, the Norse audience would know that the key to the stanza is to find another unmentioned word, one that would provide a bridge of meaning. That unmentioned word is *hæll* (heel). When substituted for the word *ekkja*, meaning "widow," it carries a double connotation; it also means "heel"—that is, "foot." The members of Egil's audience, who enjoyed the intricacies of skaldic verse, would know to replace the words *ekkjur* (widows) and *konur* (women)—both translated here as "crones"—with *hælar*, the plural of *hæll*, meaning both "feet" and "women."

Once the connection with feet is made, the rest is easy. Both nouns are connected with the adjective *allkaldar*, "thoroughly cold." Thus, the passage carries the meanings of "cold feet" and "cold women," both of which sadly afflicted Egil in his later years.

Is there a tradition of Icelandic warrior-poets complaining about women? Yes. But about their cold feet? Hardly.

Paget's Disease

In 1854 Sir James Paget became surgeon extraordinary to Queen Victoria and, a few years later, surgeon ordinary to the Prince of Wales. Paget's fame rests on his descriptions of several diseases, the most famous of which is osteitis deformans. According to Paget's classic description, this disease of the bones "begins in middle age or later, is very slow in progress, may continue for many years without influence on the general health, and may give no other trouble than those which are due to changes in shape, size, and direction of the diseased bones.... The limbs, however misshapen, remain strong and fit to support the trunk."


BETTMANN ARCHIVE

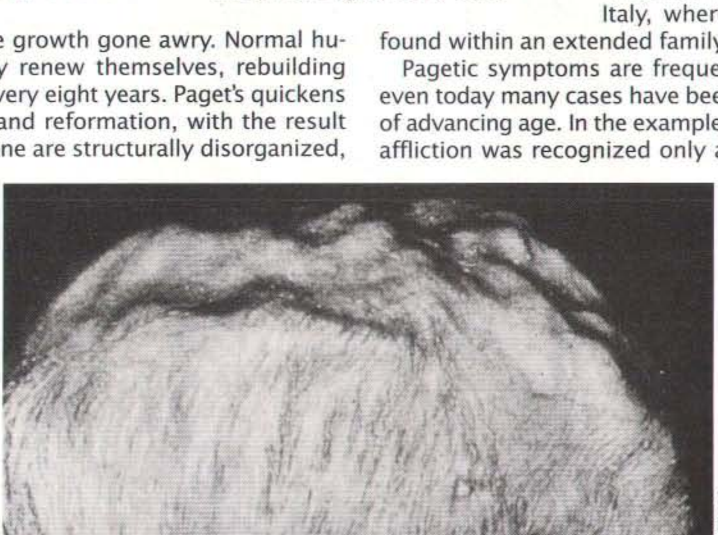
Sir James Paget, 1814–1899

Paget's disease is bone growth gone awry. Normal human bones continuously renew themselves, rebuilding completely about once every eight years. Paget's quickens the pace of breakdown and reformation, with the result that the layers of new bone are structurally disorganized, misshapen and considerably larger than the original ones. The cause of the condition could be an inherited weakness of the immune system or a virus, or both.

Paget's is an extremely old disease. The first recorded evidence of the ailment is a grossly thickened Egyptian skull dating from about 1000 B.C. It affects slightly more men than women, and it usually occurs after the age of 40.

Increasingly, statistics show that osteitis deformans is not especially rare. British studies estimate that from 3 to 5 percent of all adult males older than 40 in the U.K. have Paget's disease in some form; the frequency may reach 10 percent after age 70. In the U.S. as many as three million people probably have Paget's disease. Of these, up to 25 percent show distressing symptoms. Besides its tendency to run in families, the disease is known to cluster in specific geographical areas, mostly in Europe and particularly in England and France. Even among populations not prone to Paget's, the disease may exist, frequently limited to relatively small areas, such as the town of Avellino in Italy, where a cluster of cases was found within an extended family.

Pagetic symptoms are frequently misunderstood, and even today many cases have been attributed to the effects of advancing age. In the example of the Avellino family, the affliction was recognized only after an emigrant relative had been diagnosed in New York City. Accurate diagnosis relies on x-rays or blood tests. These examinations look for increased levels of alkaline phosphatase, a product of the cells that form bone. Urine tests may show an increased amount of hydroxyproline, another product of bone breakdown. Treatment includes drugs, specifically calcitonin and disphosphonates, which slow or block the rate of bone breakdown and formation.


MACMILLAN PRESS

CORRUGATED HEAD of a 92-year-old man is a symptom of Paget's. In the mid-1970s he was a patient of Narendra K. Chakravorty of St. Luke's Hospital in Huddersfield, England. Like Egil, his sight was affected, and he was deaf but not mentally impaired.

There is, however, a tradition of recording struggles against one's fate, including enfeeblement. In this instance the lines preserve the memory of a man's private battle with an exceptionally harrowing plight. Despite his condition, Egil still had the acuity to compose clever poetry. James Paget's classic formulation puts it this way: "Even when the skull is hugely thickened and all its bones are exceedingly altered in structure, the mind remains unaffected."

Because Egil's symptoms provide a striking picture of Paget's disease, one might ask whether the unearthing of the bones in the mid-12th century might be the source for the poetry. Could a 13th-century poet, having learned about the condition of Egil's bones, have written verses about the hardness of Egil's head using kennings? The answer is, perhaps yes, at least about the bones. Such a poet, however, would not have known the particulars of Paget's disease and could hardly have expounded on the enlarged bones to construct a detailed portrait of a man with cold feet, chills, headaches, a swaying, hanging head, inconsistent bouts of lethargy, and loss of balance, hearing and sight.

JESSE L. BYOCK

SITE OF EGIL'S BONES was the 11th-century church at Hrisbu. The ruins sit on a small hill, or church knoll. To the right is a smokehouse. The turf-built church would probably have been of somewhat similar design, though larger.

The answer is even more persuasive when we remember that the medieval text simply treats Egil's physical problems as the ravages of time. It makes no connection whatsoever between the bones and any kind of disease. In fact, the saga draws the opposite conclusion. Awed by the size and the resiliency of the skull, it points out how useful such a tough head would be for a warrior. The crucial factor is that the poetry, which may be the oldest element in the saga, independently corroborates the specifics about the bone by giving different details.

Could another disease have caused Egil's problems? I considered conditions that produce similar symptoms, such as osteitis fibrosa, acromegaly (gigantism), hyperostosis frontalis interna, fibrous dysplasia, and osteopetrosis. In each instance, however, critical symptoms do not match. By using all the sources available today, we can diagnose Egil as a probable victim of Paget's disease.

Paget's in Scandinavia

I was led to the question of Paget's disease by research intended to explain passages in a medieval saga, but it is now clear to me that the currently accepted statistics about the disease in Iceland, and possibly in all Scandinavia, are certainly inaccurate. Most studies posit an extremely low incidence, or an almost nonexistence, of the condition in this region. This situation exists because little attention has been directed to the effective diagnosis of the pathology. For example, an extensive 1982 study to determine the European distribution of Paget's disease used questionnaire replies by 4,755 radiologists. It found the disease more prevalent in Britain than in any other western European country. The study excluded Norway, Iceland, Sweden and Finland on the assumption that incidences there were very low.

Although uncommon, Paget's disease is more prevalent in modern Scandinavia than these conclusions would suggest. Until recently, the disease was thought not to exist in Iceland. During the past 10 years, however, Paget's disease has slowly but increasingly been found in modern Iceland, a fact unpublished except for a 1981 case study reported in a small journal there by Gunnar Sigurdsson of the City Hospital in Reykjavík. In July 1991, I interviewed Sigurdsson, who informed me he was treating 10 patients with Paget's disease. His observations about the symptoms closely match those of Thordur Hardarson of Iceland's National University Hospital, who was also treating patients with Paget's disease.

To the growing evidence of Paget's in Iceland, we can add the high probability that a saga-age Icelander, and perhaps even a medieval family, could have been afflicted. Recognizing these individuals as victims of Paget's disease begins to fill in the picture of the epidemiological history of the disease in early Scandinavia, providing examples of Paget's at both ends of a 1,000-year period.

Egil's poetry, Skapti's medieval observation and modern medical knowledge together provide a detailed composite of a Pagetic affliction. With this insight, we do not have to discount *Egil's Saga* to explain the misshapen skull and bones unearthed in the 12th century. On the contrary, we can see that the saga may well contain accurate information. Although we surely cannot conclude that all the sagas are historical truths, Egil's bones strongly suggest that some passages may reliably detail the past.

Is there more work to be done? Yes, for Egil's bones are possibly still buried in the old churchyard at Mosfell. We await the opportunity to unearth his remains for the third time in 1,000 years.

FURTHER READING

EGIL'S SAGA. Translated by Hermann Pálsson and Paul Edwards. Penguin Books, 1976.

CORRUGATION OF THE SKULL IN PAGET'S DISEASE OF THE BONE. N. K. Chakravorty, S. K. Das and M. S. Kataria in *Postgraduate Medical Journal,* Vol. 53, No. 615, pages 40–46; January 1977.

PAGET'S DISEASE OF THE BONE: ASSESSMENT AND MANAGEMENT. R. C. Hamdy. Praeger Scientific, 1981.

MEDIEVAL ICELAND: SOCIETY, SAGAS, AND POWER. Jesse L. Byock. University of California Press, 1988.

PAGET'S DISEASE OF THE BONE: CLINICAL ASSESSMENT, PRESENT AND FUTURE THERAPY. Edited by Frederick R. Singer and Stanley Wallach. Elsevier Science Publishing, 1991.

THE SKULL AND BONES IN EGIL'S SAGA: A VIKING, A GRAVE, AND PAGET'S DISEASE. Jesse L. Byock in *Viator,* Vol. 24, pages 23–50; 1993.

Better Than a Cure

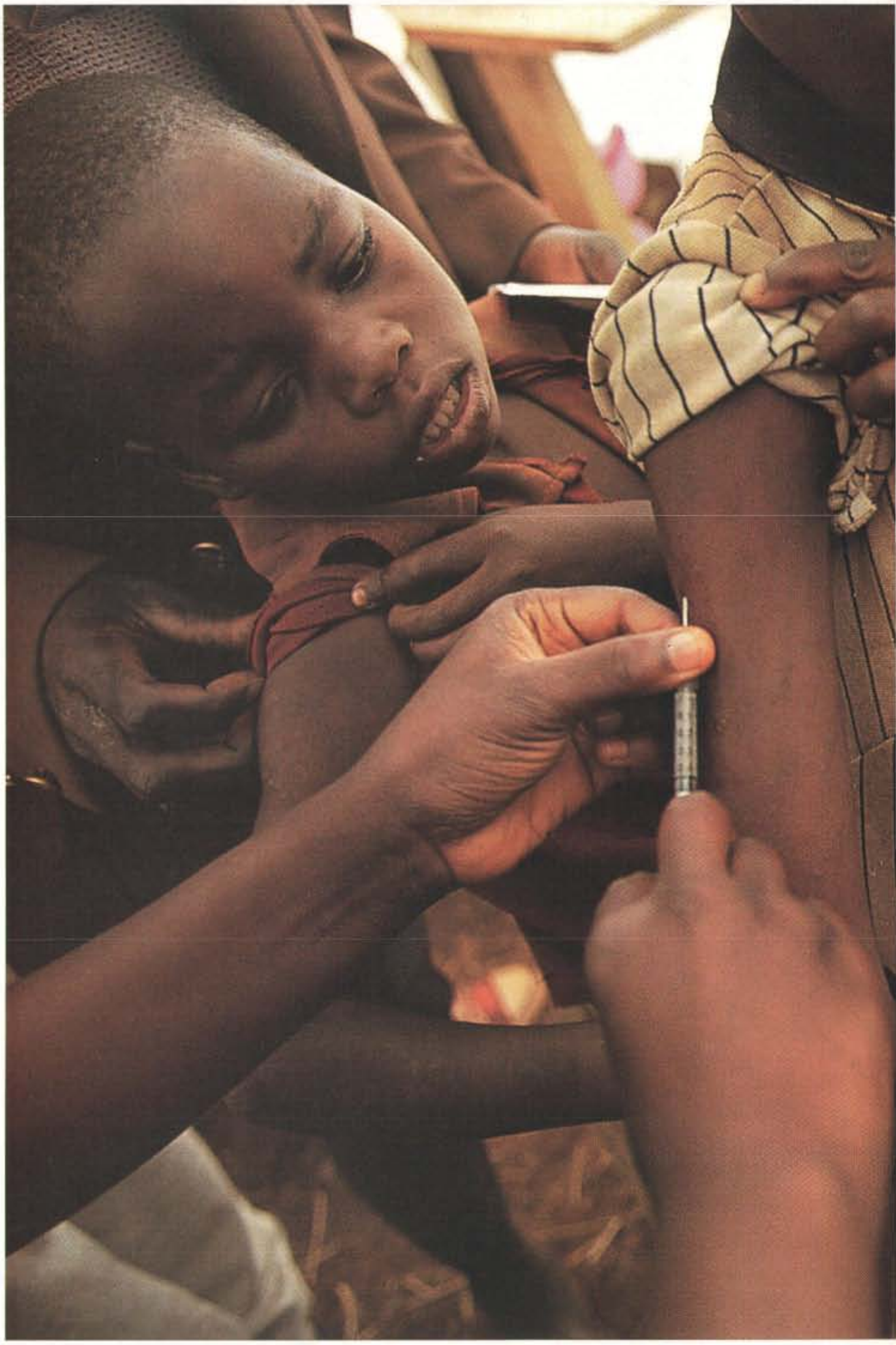
by Tim Beardsley, *staff writer*

When cholera started killing thousands of Rwandan refugees in camps in Zaire last summer, nobody was more frustrated than Jerald C. Sadoff of the Walter Reed Army Institute of Research in Washington, D.C. As chair of a World Health Organization (WHO) committee on vaccines for diarrheal diseases, Sadoff had been trying unsuccessfully for months to raise about $200,000 for research into vaccines against a newly identified strain of cholera. It took pictures of sickness and death beamed around the world from the camps to focus international attention on the disease; relief agencies eventually spent $140 million in emergency aid to contain the outbreak. "It did seem ironic," Sadoff recalls.

The episode highlights some of the difficulties faced by vaccine science. The benefits of immunization are invisible, yet governments as well as ordinary citizens are most easily persuaded to reach into their purses for medical research when they see suffering. Likewise, even though it is in the public interest for people to be well immunized, healthy individuals are unwilling to pay much for vaccines. Most pharmaceutical companies therefore avoid making them. In consequence, the global effort to develop new and more effective vaccines is small in comparison with other areas of medical research.

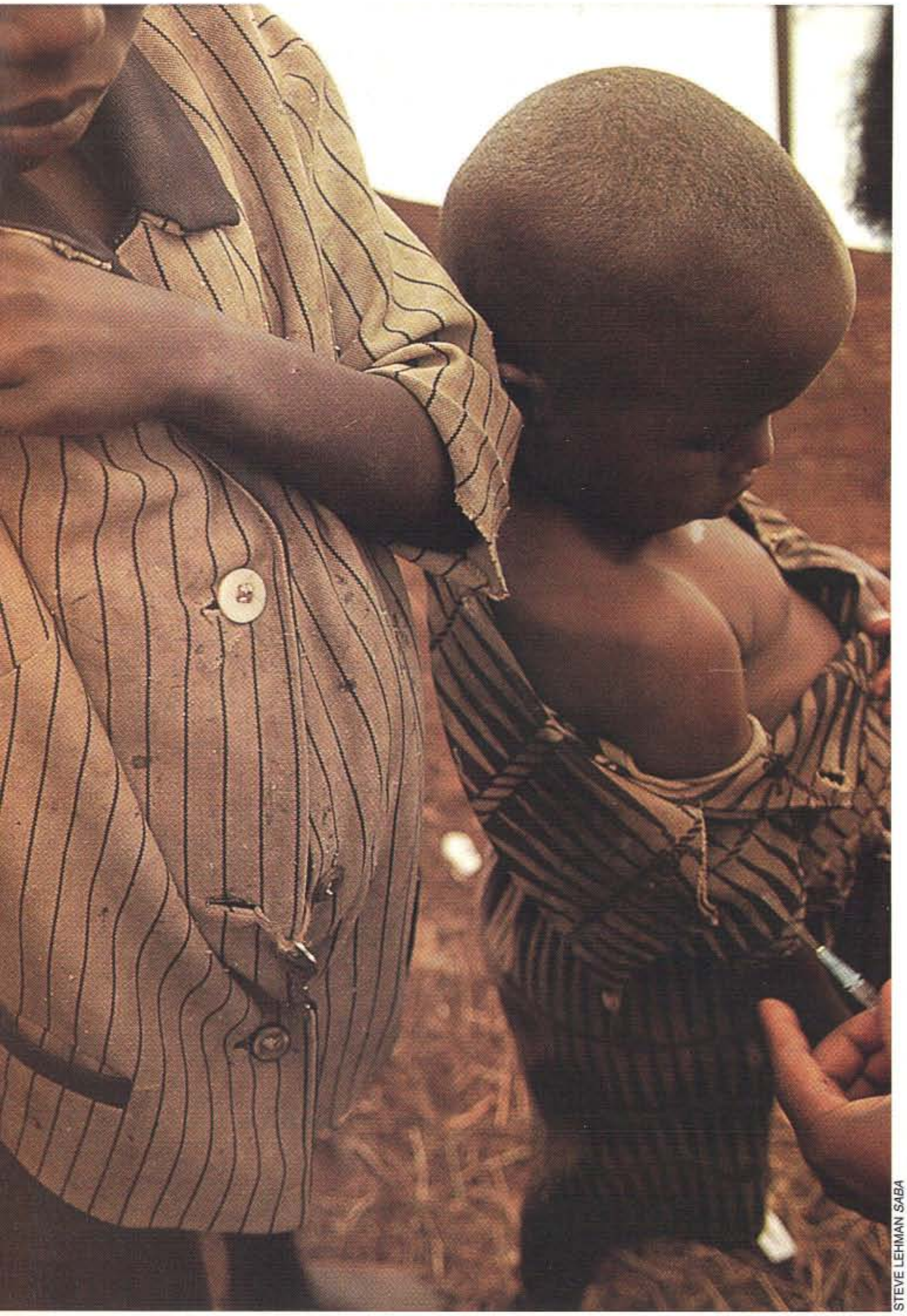
Infectious disease is still the number-one cause of death worldwide. Microbes of one kind or another kill more than 13 million every year, mostly infants. Acute respiratory infections take the biggest toll among the young; in 1992 they killed about 2.8 million children younger than five years. Diarrheal diseases took another 2.2 million, and malaria carried off a million.

Vaccines offer the best hope for reducing the appalling toll. Immunization is, simply, the best medicine. Not even antibiotics can touch it in terms of cost-effectiveness. According to David Parker and Terrel Hill of UNICEF, immunization against measles, tetanus and tuberculosis costs from $2 to $15 per "discounted year of healthy life" gained (a statistical measure of the value of a vaccine). Other common interventions cost from $25 to $1,000 for the same



VACCINATIONS against measles are administered to Rwandan refugees at a camp in Tanzania. Such immunization campaigns assume an especially high priority in situations where there is crowding and poor sanitation. Vaccines save lives and prevent disease more cost-effectively than

The World Health Organization wants industry to step up its efforts to develop new vaccines. Can big business and a public health bureaucracy see eye to eye?



STEVE LEHMAN SABA

any other form of medication. The technical means for developing new vaccines for a variety of fatal illnesses are improving steadily. Money for vaccine research remains a limiting factor, however, and often vaccines fail to reach those whom they would benefit most.

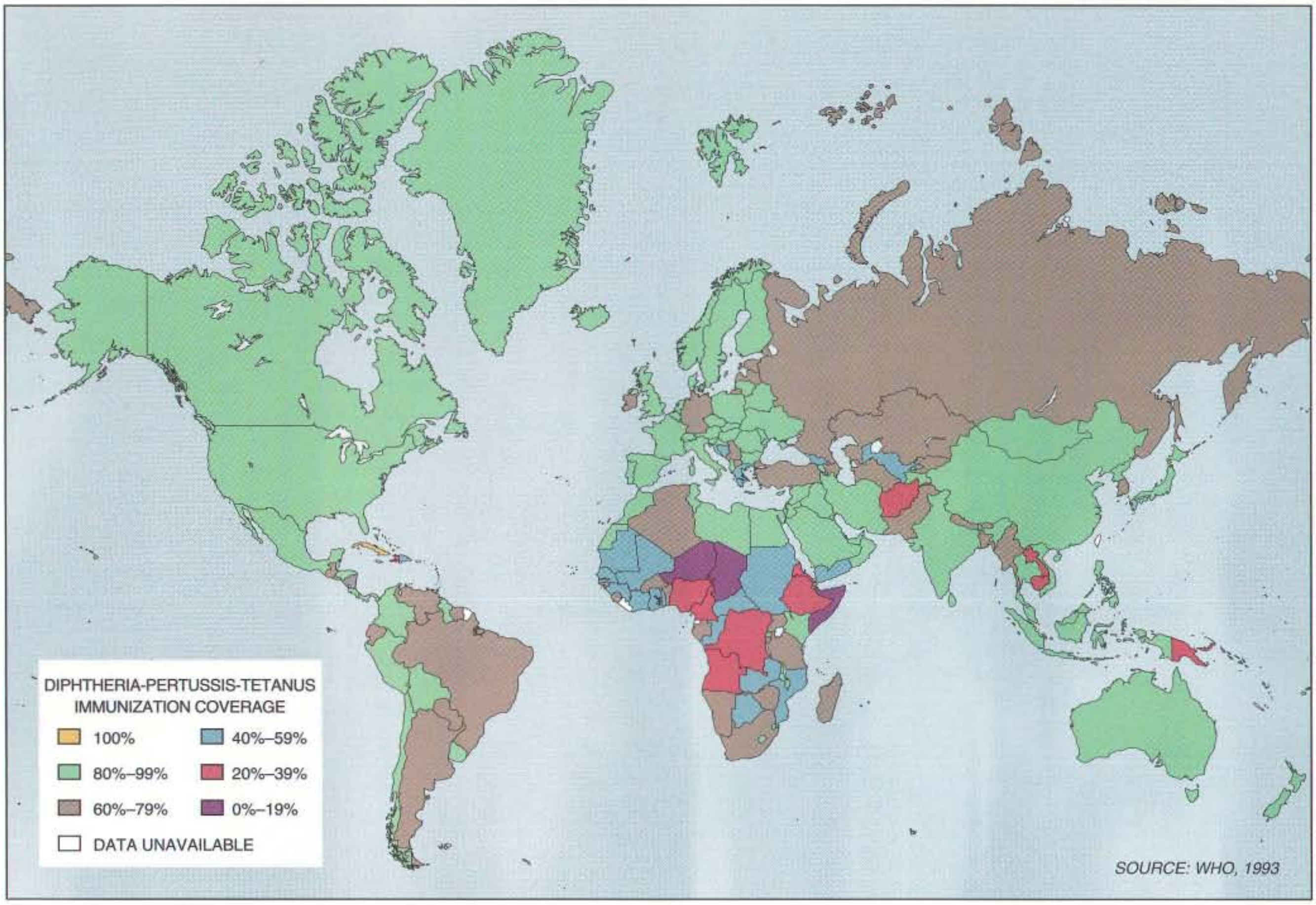
benefit. Moreover, vaccination is extremely safe. Although some serious reactions have been linked to the current generation of pertussis (whooping cough) vaccines, the risk is tiny in comparison with the threat posed by contracting the disease itself, which in the 1980s killed more than 500,000 a year worldwide. Pertussis vaccines now in clinical trials should be even safer.

Some public health experts have worried that the large reductions in death rates brought about by immunization might in poor countries be offset by children dying at greater rates from other causes, such as starvation. But the available data do not support that notion, Parker and Hill say. They even suggest that decreasing death rates from disease seem to lead to lower birth rates, presumably because women feel the need to have fewer children when more survive.

Much could be achieved merely by ensuring that today's vaccines reached all the people who could benefit from them. That step alone would save about two million lives every year, WHO calculates. Even in the U.S., an estimated 500 children die every year from diseases that could have been prevented by vaccination. Influenza, hepatitis B and pneumococcal pneumonia kill more than 500,000 adults annually in the U.S.; many of those deaths could have been avoided if available vaccines were more widely used.

There seems to be no argument among the players in the vaccine business that even larger gains against microbes are achievable with new and improved products. In the past, vaccine development was as much an art as a science. But scientific advances in recent years have pointed to fresh approaches, some of which are already proving themselves. "From an immunologic perspective, the state of knowledge is such that the art of vaccination can soon be grounded on a rational rather than an empirical basis," declares the 1993 edition of *The Jordan Report*, a survey of vaccine developments published once a year by the National Institute of Allergy and Infectious Diseases (NIAID).

At present, only about 20 different types of vaccine are in widespread use.



SOURCE: WHO, 1993

JOHNNY JOHNSON

IMMUNIZATION RATES vary widely. The map shows the proportion of infants receiving all three doses of diphtheria-pertussis-tetanus vaccine. The vaccine is one of five the World Health Organization recommends for use globally. Rates for poliomyelitis, measles and tuberculosis tend to track those shown here; for hepatitis B, coverage is more patchy.

Yet there is no shortage of leads for novel products. A survey conducted by the NIAID has identified 192 vaccine candidates that are ready for early testing in people or animals. New disease targets aside, there is plenty of room for improvement in existing products. Measles vaccine, for example, does not work in infants younger than nine months, so many children die from the illness before that age. Typhoid vaccine is unreliable. The bacillus Calmette-Guérin (BCG) vaccine against tuberculosis is only partly effective; as a result, more than three million people die of that infection every year, most of them young adults. With the spread of multidrug-resistant strains, that total seems likely to increase.

Purely practical considerations connected with delivering vaccine to its recipients are often a major hurdle. Most vaccines are sensitive to heat, so a refrigerated "cold chain" stretching from the manufacturer to the patient has to be established. The challenge can be daunting in regions where temperatures may reach more than 40 degrees Celsius and where the kerosene fuel for portable refrigerators is of varying quality and availability. The cold chain is one reason why in developing countries the cost of delivering a vaccine is far greater than the cost of the vaccine itself.

From Bench to Bush

One of the biggest problems with current products is simply that most have to be given several times to be effective. A full course of all the WHO-recommended vaccines, for example, requires five visits to a provider. Inevitably, many children do not receive all the required doses, especially in remote places. Yet although workers in the field agree that reducing the number of necessary doses is one of the most promising ways to improve protection, manufacturers have been slow to develop such vaccines, in the opinion of Philip K. Russell, a former head of the U.S. Army medical research and development command, who is now president of the Albert B. Sabin Vaccine Foundation.

Russell is not alone in believing that the progress of vaccines from the research laboratory to the field—"from bench to bush," in the argot of the business—has not been as rapid as it should be. Many vaccine researchers feel the same way. "There has existed for some time the belief that even more could be accomplished if ways could be found to bring greater cohesion" to vaccine development, the recently published *U.S. National Vaccine Plan* states.

The obstacles are complex. Manufacturers have to be convinced that the investment necessary to develop a new product—which may be anywhere from $10 million to more than $100 million—is likely to produce a good return. Vaccines, they say, are much harder than the average drug to manufacture. Some involve the use of dangerous microorganisms. "You can't imagine the number of steps it takes to make a new vaccine," says R. Gordon Douglas, Jr., president of the vaccine division at Merck & Co. To make a batch of poliomyelitis vaccine takes at least nine months, he notes.

A product is likely to generate hefty returns, however, only if it can be sold in rich countries or if the market will be extensive. Many potential disease targets are not economically justifiable. "Take Crimean hemorrhagic disease—who's going to develop a vaccine for that?" Douglas asks rhetorically. Many international development agencies, including the U.S. Agency for International Development, have been hesitant to get involved in early-stage vaccine research, Russell observes. Most agencies have seen their primary role as delivering existing vaccines to the needy. "One of the sad problems is that implementers and researchers and manufacturers hardly ever got together, and when they did they didn't talk the same language," remarks John R. La Montagne, head of vaccine research at the NIAID.

Vaccine Czar

The task of breaking that logjam has fallen to Jong Wook Lee, a soft-spoken former vaccine researcher from South Korea, who is director of the new WHO Global Program for Vaccines and Immunization. The program, which is run out of a cramped annex behind WHO headquarters in Geneva, brings the organization's vaccine research and delivery operations together for the first time. It will, Lee predicts, wield enough influence over manufacturers to persuade them to cooperate. "We are flinging open the door to industry," he announces.

Even if Lee's program meets only some of its targets for the remainder of the decade, the effort could save millions of lives. "Never in the field of human medicine will so much be owed by so many to so few," he beams in a rare immodest moment.

If Lee's ambitions seem unrealistic, it is worth remembering that WHO, in collaboration with UNICEF, has a successful history when it comes to getting vaccines to where they are needed. The organizations eradicated smallpox in 1977 and boosted global immunization levels for six major killing diseases from 5 percent in 1974 to more than 80 percent in 1990. Lee's program estimates that vaccination against just three of those six—measles, neonatal tetanus and pertussis—is now preventing almost three million deaths a year. Although immunization rates have started to flag in recent years in some countries, most of them in western and central Africa, Lee has plans to reverse that trend.

WHO has also made impressive strides against polio and estimates that vaccination prevents 550,000 cases of that crippling disease annually. WHO's American wing, the Pan American Health Organization, announced last September that an unrelenting surveillance ef-

STEVE MUREZ *Black Star*

JONG WOOK LEE is head of the new Global Program for Vaccines and Immunization at WHO and also runs the Children's Vaccine Initiative. Lee declares he is "flinging open the door to industry" in order to speed the development of new vaccines.

JOHN PINDERHUGHES

R. GORDON DOUGLAS, JR., president of Merck & Co.'s vaccine division, is angry about price controls Congress has imposed on vaccines supplied to the government. "This means we're not doing all the things we're capable of," he states.

fort combined with "vaccination days" for the entire juvenile populations of high-risk countries has already eradicated polio from the Western Hemisphere: the last recorded case there occurred in Peru in 1991. Lee's team at WHO believes polio could follow smallpox to extinction in the wild by the year 2000 (stocks of virus may exist in laboratories after then). Lee is planning to ease that passage by stepping up surveillance and vaccination efforts in countries where polio is still well known.

Close collaboration with industry on vaccine development is a new venture for WHO, however. And history suggests there are dangers. In August 1990, WHO, the World Bank, UNICEF, the U.N. Development Program and the Rockefeller Foundation launched a campaign called the Children's Vaccine Initiative (CVI). The CVI was intended to catalyze the creation of heat-stable, affordable vaccines that could be administered orally

Many Ways to Make a Vaccine

As long ago as the early 18th century, people sought protection against smallpox by pressing infectious "matter" from patients' lesions into breaks in their own skin. The highly dangerous practice was intended to cause a mild case of the disease and so bestow immunity.

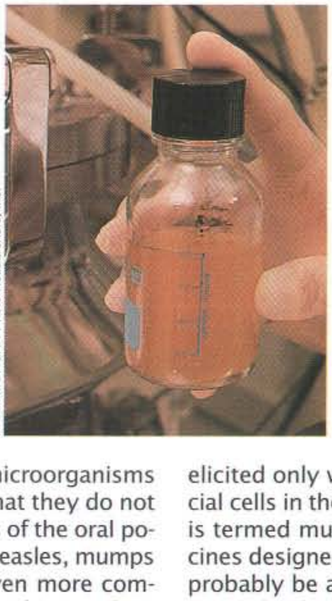
Immunization became a more reasonable proposition in 1798, when Edward Jenner demonstrated how the illness could be prevented by inoculations of cowpox, a related but less dangerous disease. The principle of using a related pathogen to provoke immunity is still being explored today in vaccines designed to protect children from rotaviruses, which cause often fatal diarrhea in millions of children. But many other approaches to vaccination are used as well.

Killed whole organisms and weakened toxins.

Several widely used vaccines, including those targeted at influenza and pertussis, are based on killed microbes. Injectable poliomyelitis vaccines also work this way. A variant of this approach, used in diphtheria and tetanus immunization, is to inject people with chemically modified versions of toxins produced by the infectious agent. Such toxoids, as they are called, allow the immune system to learn how to inactivate the poison from a real infection.

Subunit vaccines.

Vaccines consisting of molecular subunits of pathogens can avoid some of the complications of using whole organisms. Subunit vaccines are now available for meningitis, pneumonia and hepatitis B; vaccines employing this approach against respiratory syncytial virus and parainfluenza virus, both major killers, are in development. Subunit preparations are also being investigated as candidates for protection against infection with HIV (the AIDS virus) as well as malaria.



HANK MORGAN Photo Researchers, Inc.

Altered pathogens.

Some diseases require a more powerful immune stimulus. In such cases, favorable results can sometimes be obtained with live microorganisms that have been weakened, or attenuated, so that they do not produce significant illness. This is still the basis of the oral polio vaccine and the combined vaccine against measles, mumps and rubella. Modified pathogens might be even more commonly used in the future. Genetic engineering has made it possible to introduce immune-stimulating proteins from a range of pathogens into tried and trusted carrier organisms such as vaccinia (derived from the cowpox virus with which Jenner countered smallpox) and various bacteria.

Conjugates.

Vaccines for some bacterial diseases, such as pneumococcal pneumonia and meningitis, cannot be used in babies, because their immune systems do not recognize as foreign the sugars in the bacterial cell walls. During the past two decades, however, researchers have learned how to combine these sugars with protein carriers. The resulting "conjugate" vaccines work well even in infants. In 1986 the first conjugate vaccine, against *Hemophilus influenzae* type B, was licensed in the U.S., and current versions are effective in children as young as two months. Conjugate vaccines for pneumococcal pneumonia and meningococci groups A and C are under development. "They could have a tremendous impact in a number of diseases," says John R. La Montagne of the National Institute of Allergy and Infectious Diseases.

Adjuvants and microspheres.

Many improvements in vaccines expected over the next decade are likely to be the result of better adjuvants or carriers. Adjuvants are substances that potentiate an immune response. The only adjuvant licensed for use in humans at present is alum, an aluminum salt, but other chemicals, including some complex organic ones, are being studied. Among them are muramyl dipeptide, squalene, lipid spheres known as liposomes and cagelike organic structures called immunostimulatory complexes (ISCOMs).

An approach that has long captured the imagination of researchers puts immunogenic chemical fragments into minute polymer spheres that only slowly let the immune system "see" their contents as they diffuse out. The scheme might be able to confer long-lasting immunity for some diseases with a single inoculation. Recently workers have successfully stabilized tetanus toxoid in polyester microspheres, thus pointing the way to a one-dose tetanus vaccine.

Oral vaccines.

Most vaccines are thought to exert their effects in the bloodstream. Oral polio vaccine, however, seems to work by triggering a different kind of immune response that can be elicited only when immunostimulatory molecules reach special cells in the lining of the gut. Immunity generated this way is termed mucosal immunity. Many researchers believe vaccines designed to stimulate mucosal immunity—which would probably be administered orally—might protect against diseases that have so far proved resistant to vaccination. Sexually transmitted diseases, including HIV infection, are a major focus of research interest, as are pathogens that enter the body through the gut, such as *Vibrio cholerae* (which causes cholera) and *Shigella* (dysentery).

Naked DNA.

The most recent and surprising development was the demonstration that DNA, when injected into muscle, can by itself early in life and would protect against a wide range of diseases.

The plan was to cajole manufacturers into making innovative products by establishing priorities for research, organizing clinical trials and raising money. Ciro A. de Quadros of the Pan American Health Organization says the initiative has secured international agreement on procedures for manufacturing vaccines and checking their quality, an important step toward being able to make novel combinations. It also stimulated a promising project that demonstrated how deuterium oxide, better known as heavy water, can be used to make a polio vaccine that is stable for up to a week at 37 degrees C. (The product is now in industrial development.)

Russell believes that the CVI's greatest achievement was to persuade manufacturers to make vaccines that combine protection against hepatitis B and *Hemophilus influenzae* type B (Hib)—a common and dangerous cause of meningitis—with the existing combination vaccines for diphtheria, pertussis and tetanus. "There was tremendous embedded resistance" from people who saw no need to disturb the status quo, Russell recalls. "If you look at the CVI as a global intellectual movement, it's been a substantial success."

Others have a different assessment. "European industry was quite upset with the CVI because it ignored our achievements," explains Walter S. Vandersmissen, director of government affairs for SmithKline Beecham in Brussels. "It was a dream, and we tried to instill a sense of realism. Today not much has been achieved." The ambitions were slowly scaled back, but friction still developed between the CVI and WHO when workers for those agencies started giving conflicting advice to governments. Moreover, the CVI failed to raise more than a few million dollars toward a planned $300-million fund for vaccine development.

Mounting disagreements among the CVI's sponsors over what the initiative should be doing were finally resolved last August. The sponsors agreed to let the initiative be advised by the same committee that oversees WHO's new program, thus effectively bringing them under one command. Lee was given the additional appointment of executive secretary of the CVI, which makes him almost a global vaccine czar. The CVI will now concentrate on raising awareness of research results and gaining political commitments to improved vaccines, according to Roy Widdus, an adviser to Lee on the project.

Whether the new WHO effort can forge a consensus on vaccine development will depend on the ability of Lee and his team to better the record of the CVI in building constructive links with industry. Lee acknowledges that WHO has in the past been coy, even "a little paranoid," about close relations with manufacturers, fearing that its impartiality might seem to be tainted. But he now has good reason to want to change that stance.

Last year WHO received the results of a confidential study conducted by Mercer Management Consulting on the economics of the $3-billion global vaccine industry. A central finding was that the $60 million that UNICEF spends ev-

confer immunity. Workers at Vical, a biotechnology company in San Diego, Calif., collaborating with Merck and U.S. Navy investigators, have shown that such "naked DNA" can immunize mice against malaria and influenza. The DNA, which encodes a protein displayed by the pathogen, apparently stimulates host tissues to synthesize proteins that the immune system recognizes as foreign.

By continuously stimulating the immune system, naked DNA vaccines could produce responses as strong as those induced by attenuated organisms. In particular, they seem able to stimulate the arm of the immune system that employs *T* cells to kill invaders. Naked DNA technology "has an immensely powerful capability," says Philip K. Russell of the Albert B. Sabin Vaccine Foundation.

SOME VACCINES IN DEVELOPMENT

Disease/Pathogen	Technology
Respiratory syncytial virus	Attenuated virus; subunit in microspheres
Influenza	Attenuated virus and naked DNA
Group B streptococci	Conjugates
Parainfluenza	Subunits and attenuated virus; inactivated virus in microspheres
Meningococci group B	Modified polysaccharide
Measles	Subunits in ISCOMs and in vaccinia; attenuated virus; naked DNA
Pneumococcal pneumonia	Subunits and conjugates
Typhoid	Subunits and conjugates
Cholera	Inactivated and attenuated pathogens; subunit combinations
Shigella	Bacterial vector; subunits and conjugates
Tuberculosis	Mycobacterial vector; subunit
Malaria	Subunit and naked DNA
Dengue	Yeast, yellow fever and vaccinia vectors; attenuated and chimeric viruses
Rotavirus	Attenuated and modified viruses
Tetanus	Single dose: toxoid in microspheres
Schistosomiasis	Subunit
Human immunodeficiency virus	Subunits; vaccinia vectors; inactivated virus

SOURCES: Global Program for Vaccines and Immunization, World Health Organization;
The Jordan Report, *National Institute of Allergy and Infectious Diseases*

ery year to buy vaccines for poor countries—at discount prices—was enough for the deals to be profitable for manufacturers, provided they could charge higher prices elsewhere. The study thus made clear that UNICEF—which acts as the purchasing arm of WHO—could exert pressure on manufacturers. "We are writing the guidelines, so we have leverage," Lee declares. "No one country can perform this function, and industry cannot ignore the process."

Building Bridges

Lee is keen to stimulate vaccine production in developing countries as well as in the industrial world. About 60 percent of the global supply of diphtheria-pertussis-tetanus vaccine—a relatively simple product—is made in developing countries; Lee explains that he would be quite happy to see manufacturers outside the industrial world grab a larger slice of the pie for other products. "In the developing world, they are not just driven by profit motivations," he says.

At present, technical problems mean some vaccines manufactured in developing countries "may be completely useless," Lee admits. But WHO has introduced standards of good manufacturing practice that should eliminate such variations. Lee argues that if manufacturers in Asian countries can make cars suitable for the U.S. market, Asian pharmaceutical companies can probably make vaccines to WHO's standards.

Lee is also spearheading an effort to encourage more countries to make or buy vaccines rather than appeal to the charity of UNICEF. At present, UNICEF purchases fully 40 percent of the world's pediatric vaccines for countries that have not made their own arrangements. Some of those countries could buy for themselves, Lee says, and a few—including China and India—will be told to start doing so immediately. Others will be given a grace period. The move is likely to be well received by manufacturers: because many poor countries are unlikely to establish their own plants, Lee's policy means more vaccines will be sold at higher prices.

Vaccine makers have noticed the new industry-friendly stance in Geneva. "It's a major change for a group that has a tendency to denigrate the profit-making sector," reflects Thomas M. Vernon, Jr., executive director for medical, scientific and public health affairs at Merck. "There has been a wind of change," Vandersmissen agrees. "Now industry can frankly state its point of view, including financial concerns."

Within the past year the Global Program for Vaccines and Immunization has awarded $400,000 to industrial and academic researchers to investigate how a heat-stable oral polio vaccine might be manufactured. It has also entered into contracts worth more than $300,000 with companies in the U.S. and Europe to develop improved vaccines for group A and C meningococci and for tuberculosis and leprosy, as well as a single-dose tetanus toxoid using microsphere technology [*see box on pages 72 and 73*]. The amounts are small, but manufacturers can "feel they are partners," says Paul-Henri Lambert, head of the vaccine R&D division.

U.S. vaccine producers, however, are unlikely to become suppliers for UNICEF. According to Ronald J. Saldarini, president of Lederle-Praxis Biologicals, another major U.S. vaccine manufacturer, the prices are simply too low. "I can't afford to supply" UNICEF, he states. Merck's Douglas blames the U.S. government for making it hard for domestic. manufacturers to bid on UNICEF contracts. Merck would do so, he says, if it could charge lower prices than it charges the federal government for vaccines. But back in 1982, Merck executives were lambasted by a congressional committee for proposing to do just that, and the company wants to avoid a repeat episode. Overseas manufacturers that have supplied UNICEF programs were encouraged to do so by their governments as a matter of foreign policy, Douglas maintains.

Constraints on vaccine prices also lie at the heart of a separate bitter argument U.S. manufacturers have with their government. The amount of vaccine that U.S. producers sell to the Public Health Service can fluctuate drastically from year to year as a consequence of the rules used to apportion purchases. Yet under the Vaccines for Children program [*see box on opposite page*], which started last October, the price the government pays is capped, with increases only for inflation. The combination is anathema to industry. Barry R. Bloom, a researcher at the Albert Einstein College of Medicine in Bronx, N.Y., argued recently in *Science* for the creation of a national vaccine commission to try to bring some stability to the rocky relations between the U.S. government and its suppliers. The commission would assess vaccine supply and demand and be able to respond to emergencies, such as an outbreak of infectious disease. "If there were an outbreak of yellow fever in New Orleans, we would be out of vaccine in two weeks," Bloom declares. "We are not prepared."

Others are not convinced that a commission would help. Saldarini of Lederle-Praxis sees no need for another government committee. "I'm committeed out," he complains. In any event, Congress recently directed the nearest thing the government had to a central authority on vaccines—the National Vaccine Program Office—to cut 30 people from its staff of 35. Russell calls the effective dismantling of the office "a disaster."

Lee seems anxious to avoid similar disasters in the international arena. But

PETER CHARLESWORTH *SABA*

MOBILE VACCINATION CLINIC near Chang Rai, Thailand, is visited by hill tribe villagers. In such remote areas lacking electricity and modern refrigeration, distribution efforts are complicated by the need to keep the vaccines cool.

Born in the U.S.A.

Even in a rich country such as the U.S., getting vaccines to infants is difficult. Almost all American children are fully vaccinated by the time they enter school, since vaccination is a condition for enrollment. Still, many do not receive all their shots by the recommended age. Henry D. Mustin and his colleagues at the University of Washington published in a recent issue of the *Journal of the American Medical Association* a study showing that in a nationally representative sample only 46 percent of white infants and 34 percent of black infants had received adequate immunizations by the age of eight months. According to the Centers for Disease Control and Prevention, 67 percent of U.S. youngsters have received all the recommended doses of diphtheria-pertussis-tetanus, measles-mumps-rubella and poliomyelitis vaccines by the time they are two years old. But only 55 percent had received the three recommended *Hemophilus influenzae* type B shots, and only 16 percent had been properly vaccinated against hepatitis B. The vaccine-policy joke (there is only one) is that there are counties in Texas where the cattle are better immunized than the children are.

Missed vaccinations can have serious consequences. Poor measles vaccination rates resulted in a 55,000-case outbreak in the U.S. in 1989–1991, causing 136 deaths. Concern about the low rates prompted the government to implement last year the Vaccines for Children (VFC) program, which was intended to make it easier for uninsured or inadequately insured children to obtain free vaccines from private physicians. (Private pediatricians typically charge $270 for a full course of 11 shots, although vaccinations have also been free for the asking in public health clinics.) But the program has gotten off to a shaky start.

Initially the government had proposed distributing vaccines for the program to private physicians from a single warehouse in Burlington, N.J. Officials realized belatedly, however, that they could not establish the necessary tracking and handling system by the program's October 1 starting date; consequently, that plan was scrapped. Some states with experience in vaccine distribution have agreed to disburse VFC vaccines to physicians themselves, but as of early November, 24 states had still not done so.

U.S. vaccine manufacturers bitterly oppose the VFC program because it has increased the fraction of domestic vaccine bought by the government (at a half-price discount) from about 50 to 80 percent. At the same time, Congress put a cap on prices, so makers cannot compensate for the fall in revenues. The result, they say, will be less research and fewer new products.

R. Gordon Douglas, Jr., of Merck notes that the projected shortfall is already constraining the company's choices about which projects to pursue. "These decisions will affect health 10 years from now, and I can only tell you that this means we're not doing all the things we are capable of," he insists. Christine M. Grant of Connaught Laboratories, which is owned by Pasteur-Merieux, remarks that the price-capping provisions of the program "send a message that is 180 degrees opposed to support for research and development." Ronald J. Saldarini, president of Lederle-Praxis Biologicals, charges that the VFC program "is an attempt to virtually nationalize the industry and take us out of the private sector."

Even if the VFC program is fully implemented, Douglas argues that it will not significantly increase the number of children in the U.S. who receive their immunizations on time. The reason some children fall behind is not cost, Douglas maintains, but a lack of education and ready opportunities in impoverished inner-city areas. Douglas's proposed solution to U.S. vaccination woes is to require that every health insurance package cover preventive services for children.

GENERAL SERVICES ADMINISTRATION

The VFC vaccine distribution center that wasn't

whether the new WHO/CVI combined bureaucracy can really spur the development of more effective vaccines is still an open question. Despite Lee's overtures to industry, difficulties remain. Lambert says industry's penchant for secrecy about its plans—especially when it scents profits—is still a problem. Vandersmissen argues that WHO would be most effective if it concentrated on organizing clinical trials.

Lee has his own measure of progress in relations with industry. "When we invite, they always come," he observes with satisfaction. "They are never too busy." But in the long term, success will depend on other factors, not least on how much money the program can raise. "Unless some other source of funds is found, none of this can really work," Bloom says. The Global Program for Vaccines and Immunization now has a budget of less than $30 million; Lee would like to increase that to $50 million.

More important still will be whether WHO adopts economically sustainable policies that will keep the industry connection strong. It is still too early to judge the program, but Lee at least appears to have a clear vision. "Whatever else," he admonishes, "we must not kill the goose that lays the golden eggs."

FURTHER READING

THE JORDAN REPORT: ACCELERATED DEVELOPMENT OF VACCINES IN 1993. National Institute of Allergy and Infectious Diseases, National Institutes of Health, 1993.

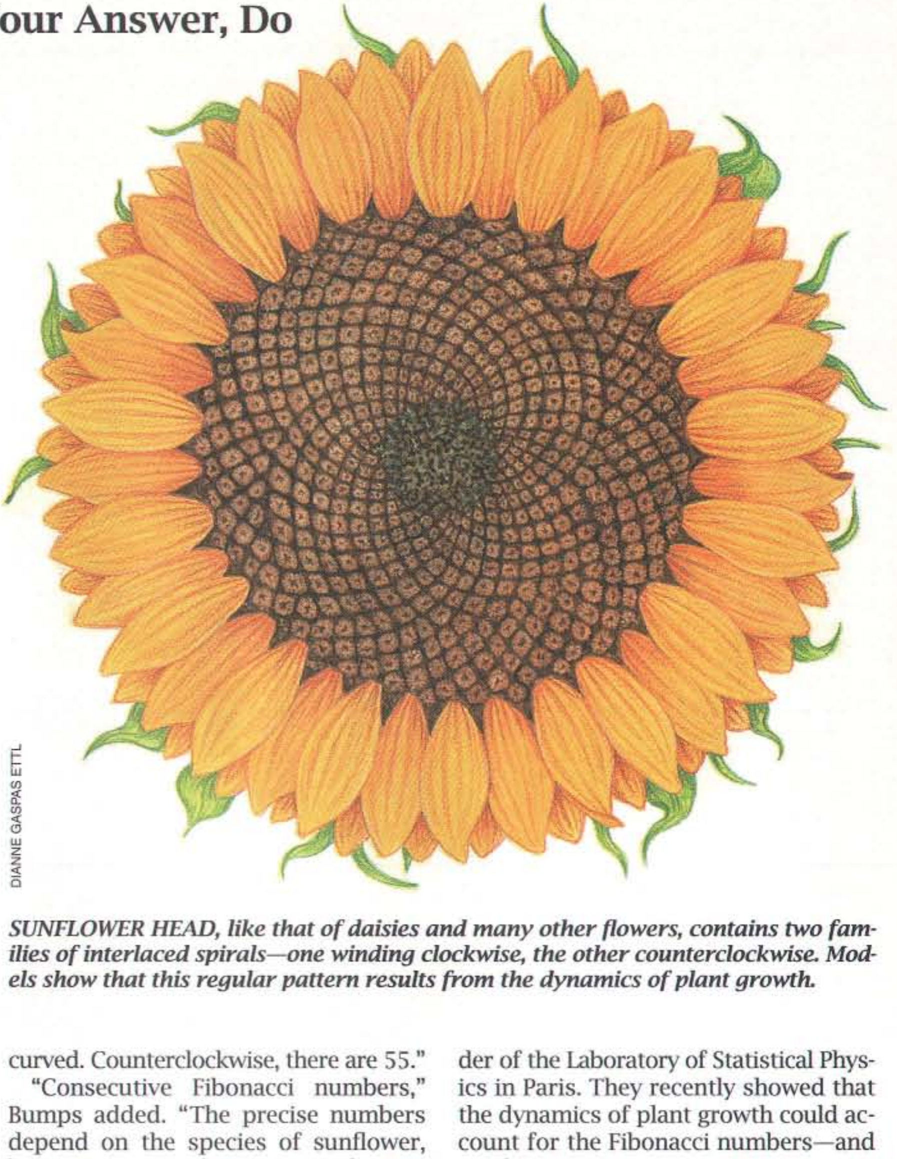
FRONTIERS IN MEDICINE: VACCINES. Special section in *Science*, Vol. 265, pages 1371–1404; September 2, 1994.

VACCINES AND PUBLIC HEALTH: ASSESSING TECHNOLOGIES AND PUBLIC POLICIES. Special section in *International Journal of Technology Assessment in Health Care*, Vol. 10, No. 1, pages 1–196; Winter 1994.

MATHEMATICAL RECREATIONS

by Ian Stewart

Daisy, Daisy, Give Me Your Answer, Do



DIANNE GASPAS ETTL

SUNFLOWER HEAD, like that of daisies and many other flowers, contains two families of interlaced spirals—one winding clockwise, the other counterclockwise. Models show that this regular pattern results from the dynamics of plant growth.

Bees buzzed drowsily, the sun beamed down and sunflowers waved in the breeze. Grimes the shepherd boy snored under a tree, while Bumps the goose girl made daisy chains. Suddenly she stopped.

"Grimes! I've just found a daisy with 31 petals. Usually this kind has 34."

Grimes sat up and stretched. "Really? It's curious that there's a specific number. Though I suppose that the flower's genes must specify—"

"I don't see that they must. I mean, genes tell plants how to make chlorophyll, but they don't tell them to make it green. That's chemistry, not genetics."

Grimes had been through this argument with her before. "Yeah, sure. Some features of the morphology of living creatures are genetic in origin, and others are a consequence of physics, chemistry and the dynamics of growth."

"Right," Bumps said. "Genetics can give rise to pretty much anything, whereas physics, chemistry and dynamics produce mathematical regularities."

"I wouldn't say that 34 is a very striking regularity," Grimes said.

Bumps pulled petals from her daisy. "Agreed, but the numbers that arise in plants—not just for petals but for all sorts of other features—are normally very special. Lilies have three petals, buttercups have five, marigolds 13, asters 21, and most daisies have 34, 55 or 89. You don't find any other numbers very often. The main exceptions are when those same numbers occur doubled or when the so-called anomalous series appears—3, 4, 7, 11, 18 and so on."

Grimes scratched his head. "I've seen those numbers before."

"Yes. The numbers 3, 5, 8, 13, 21, 34, 55, 89 form the beginning of—"

"The Fibonacci series," Grimes said in triumph. "Each number is the sum of the two that precede it. Your 'anomalous series' has the same pattern."

"Right," Bumps replied. "Fibonacci was a medieval mathematician who invented the series to model the growth of rabbit populations. That didn't work very well. But his numbers turn up in many different places. Look at how a sunflower's florets are arranged."

Grimes eyed a nearby plant myopically. "Wow. Spirals—34 spirals wind clockwise, like the spokes of a wheel, but curved. Counterclockwise, there are 55."

"Consecutive Fibonacci numbers," Bumps added. "The precise numbers depend on the species of sunflower, but you get 34 and 55, or 55 and 89, or even 89 and 144. Daisies, too."

"Weird," Grimes said.

"Entirely," Bumps agreed. "If genetics can give a flower any number of petals it likes, why such a preponderance of Fibonacci numbers?"

Grimes snapped his fingers. "You're telling me that the numbers arise through some mathematical mechanism? Physics, or chemistry, or—"

"Dynamics," Bumps said firmly.

"Has somebody actually explained how plant growth might yield Fibonacci numbers?" Grimes asked.

"Well, lots of people have suggested many different kinds of answers. But for me, the most dramatic insight comes from Stéphane Douady and Yves Couder of the Laboratory of Statistical Physics in Paris. They recently showed that the dynamics of plant growth could account for the Fibonacci numbers—and much more.

"The basic idea is an old one," Bumps continued. "If you look at the tip of the shoot of a growing plant, you can detect the pieces from which all the main features of the plant—leaves, petals, sepals, florets or whatever—form. At the center of the tip is a circular region of tissue having no special features, called the apex. Around the apex, one by one, tiny lumps called primordia emerge. Each primordium migrates away from the apex and eventually develops into a leaf, petal or the like. So you must explain why you see spiral shapes and Fibonacci numbers in the primordia."

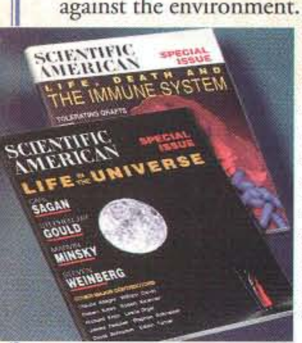
"How?" Grimes wondered.

"The first step is to appreciate that the spirals most apparent to the human

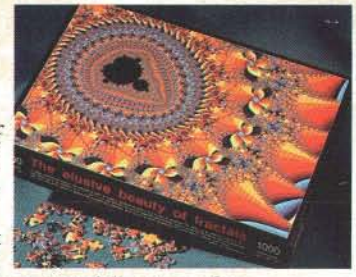
SCIENTIFIC AMERICAN SELECTIONS

YOUR PASSPORT TO INTELLECTUAL ADVENTURE

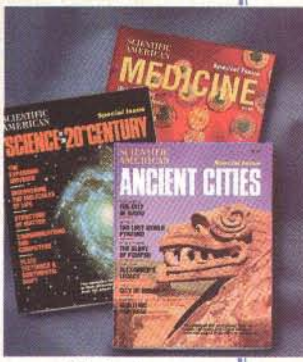
OUR ANNUAL SINGLE-TOPIC ISSUES are always in demand. Choose 1994's, ***Life in the Universe,*** and explore our universe's past and prospects. Or for 1993, ***Life, Death and the Immune System,*** and investigate the human body's single wall of defense against the environment. Each special issue contains a series of contemporary articles by top scientists and authors examining its subject in depth. Each is just $6.00, including s&h.



PREPARE YOURSELF for the elusive beauty of **Spektrum's fractal puzzle.** A single mathematical formula generates a multitude of patterns and variations in this 1000-piece, 19.6" x 27.5" puzzle, based on Benoit Mandelbrot's set for the function, $f(z)=z \tan z+c$, which includes the famous "apple" pattern. But get this dazzling full-color puzzle just for the fun of it. $29.95 plus $3.95 s&h; $29.00 for each additional puzzle.

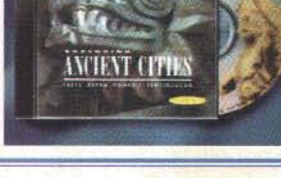


SPECIAL ISSUES are not included with annual subscriptions. ***Ancient Cities*** explores early urban planning. Learn what secrets key sites have revealed about our past and future. ***Medicine*** sheds new light on heart disease, cancer, AIDS, the immune system, illicit drugs, the future, and more. ***Science in the 20th Century*** includes plate tectonics, the expanding universe, the character of matter, the chemistry of life, and computer communications. Each title is just $6.00, including s&h.

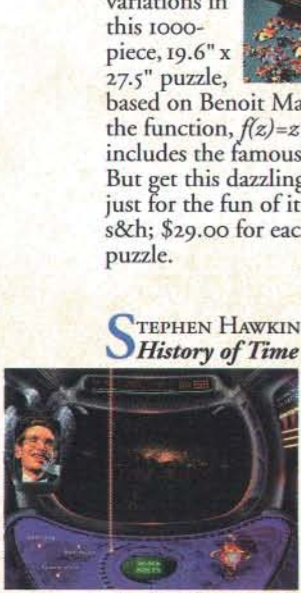


EXPERIENCE THE ANCIENT WORLD with ***Exploring Ancient Cities,*** an interactive CD-ROM, for both Mac and Windows. Tour four spectacular ancient cities and vanished civilizations with video, still photography and sound effects. Articles from SCIENTIFIC AMERICAN's anthology, *Ancient Cities,* are included explaining how these cities flourished and why they did not survive. $59.95 plus $5.00 s&h.

This CD-ROM is in a hybrid format. **Minimum system requirements:** A CD-ROM drive with a sustained transfer rate of 150 Kbps. *Macintosh:* System 7.0; 6MB RAM; 13" monitor displaying 256 colors. *IBM compatibles:* 486/33MHz CPU; 8MB RAM; DOS 5.0; Windows 3.1; VGA+ monitor displaying 256 colors at 640x480 resolution; sound card.

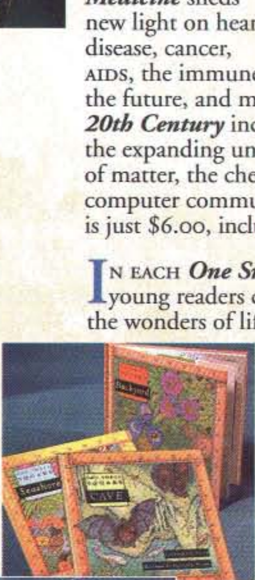


STEPHEN HAWKING'S ***A Brief History of Time*** Mac or Windows CD-ROM is an interactive journey to distant galaxies, black holes, and alternate dimensions through extraordinary graphics and narrative. Centuries of scientific thought and supreme questions about the origin and nature of the universe are explored. The book's text is included. $59.95 plus $5.00 s&h.



Minimum system requirements: *A Brief History of Time* is a CD-ROM available in both Macintosh and Windows formats and requires a CD-ROM drive with a sustained transfer rate of 150Kbps (double-speed recommended). *Macintosh:* System 7.0; 6MB RAM; 13-inch monitor with 256 colors. *IBM compatibles:* 486/33MHz CPU; 8MB RAM; DOS 5.0; Windows 3.1; VGA+ monitor displaying 256 colors at 640x480 resolution; sound card.

IN EACH ***One Small Square*** BOOK, young readers can intimately explore the wonders of life and ecology with the guidance of these beautifully illustrated books. Choose ***Backyard, Seashore*** or the latest title, ***Cave,*** for $14.95 plus $1.95 s&h. Order two or more for just $14.00 each. While the kids explore, pick up a deck of our own SCIENTIFIC AMERICAN **playing cards,** decorated with 19th century art from our archives. $7.95 per standard deck plus $1.95 s&h.



HERE'S MY MANIFEST. START MY ADVENTURE.

Name *(print clearly)*

Address Apt.

City State Zip

Method of payment (no cash or c.o.d.'s, please):

☐Check ☐Money order

☐Visa ☐MasterCard ☐American Express

Card number Expiration date

Signature

☐Please send me a free catalog. RP195

Item	Quantity	Price each	Price each 2 or more	Item sub-total	S&H each	Total
Life/Immune System		$6.00	—		Free	
Life In The Universe		$6.00	—		Free	
Fractals puzzle		$29.95	$29.00		$3.95	
Ancient Cities		$6.00	—		Free	
Medicine		$6.00	—		Free	
Science/20th Century		$6.00	—		Free	
Exploring Anc't Cities		$59.95	—		$5.00	
Brief History of Time		$59.95	—		$5.00	
Backyard Small Square		$14.95	$14.00		$1.95	
Seashore Small Square		$14.95	(mix or		$1.95	
Cave Small Square		$14.95	match)		$1.95	
Playing cards		$7.95	—		$1.95	

Send your order to: Scientific American Selections
PO Box 11314; Des Moines, IA 50340-1314

Total order:

Sales tax:*

Grand total:

*Residents of CA, IA, IL, MA, MI, NY please add appropriate sales tax. Canadian residents please include GST. GST no. R127387652.

OR CALL 1-800-777-0444 AND START YOUR ADVENTURE NOW!

PLACE YOUR ORDER TODAY!

eye—the parastichies—are not fundamental. The most important spiral is formed by considering the primordia in their order of appearance [*see illustration on this page*]. Primordia that appear earlier travel farther, so you can deduce their order based on their distance from the apex. You find that the primordia are spaced rather sparsely along a tightly wound spiral, called the generative spiral. With me so far, Grimes?"

"No problem. But why do the primordia form spirals?"

"That comes a bit later. The pioneer crystallographer Auguste Bravais and his brother, Louis, observed one essential quantitative feature in 1837. They drew lines from the center of each primordium to the center of the apex and measured the angles between successive primordia, as seen from the center of the apex. Look at the angle between the primordia numbered 29 and 30, or 30 and 31. What do you notice?"

Grimes squinted. "They look the same."

"Precisely. The successive angles are pretty much equal; their common value is called the divergence angle. The primordia are equally spaced, in an angular sense, along the generative spiral. How big do you think the divergence angle is?"

"Quite big, more than a right angle."

"Good. It's usually close to 137.5 degrees." Bumps looked smug, although Grimes didn't know why.

"Take consecutive numbers in the Fibonacci series," Bumps began to explain.

"Like 34 and 55?"

"Exactly. Now, form the fraction 34/55 and multiply by 360 degrees."

Grimes fished out the pocket calculator he normally used to keep track of feedstocks. "Um. It's 222.5 and a bit."

"You can measure angles externally or internally. Your answer is more than 180 degrees, so subtract it from 360."

"Right," Grimes said, punching the buttons. "That's 137.5 degrees."

"You got it. The ratio of successive Fibonacci numbers, as they get bigger, gets closer to 0.618034, which is $(\sqrt{5}-1)/2$, the so-called golden number, denoted by the Greek letter phi, φ."

"I thought the golden number was $(\sqrt{5}+1)/2$," Grimes queried.

"That's 1.618034. The golden number equals both $1+\varphi$ and $1/\varphi$. If you look at the ratio the other way up, say, $55/34 = 1.6176$, then the limit of increasingly larger Fibonacci ratios tends toward 1.618034 instead. At any rate, the key to the whole shebang is the 'golden angle,' which is $360(1-\varphi)$ degrees, or 137.50776 degrees. The Bravais brothers observed that the angle between successive primordia is very close to the golden angle."

"Gotcha."

"If you plot successive points on a tightly wound spiral at angles of 137.5 degrees, because of the way neighboring points align, you get two families of interpenetrating spirals. And because of the relation between Fibonacci numbers and the golden number, the numbers of spirals in the two families are consecutive Fibonacci numbers."

Grimes watched the butterflies flitting about the meadow for a moment. "So it all boils down to explaining why successive primordia are separated by the golden angle?"

"Yes. Everything else follows from that, provided you assume that successive primordia spring up around the edge of the apex—as suggested by Wilhelm Hofmeister in 1868—and that they move away in a radial direction."

"Does the speed at which they move matter?" Grimes asked.

"Definitely. Because of the way plants grow, migrating primordia actually speed up as the radius increases—keeping a velocity proportional to the radius."

"And that's where this theory of Douady and Couder comes in?"

"That's right," Bumps said. "They built their ideas on an earlier insight. If you model plant seeds as circular disks having a fixed radius and pack them together as closely as possible, while retaining a constant divergence angle of 137.5 degrees, then the *n*th seed (counting from the newest to the oldest) must be placed at a distance proportional to the square root of *n*. So the golden angle allows the seeds to pack most efficiently."

"Say again?"

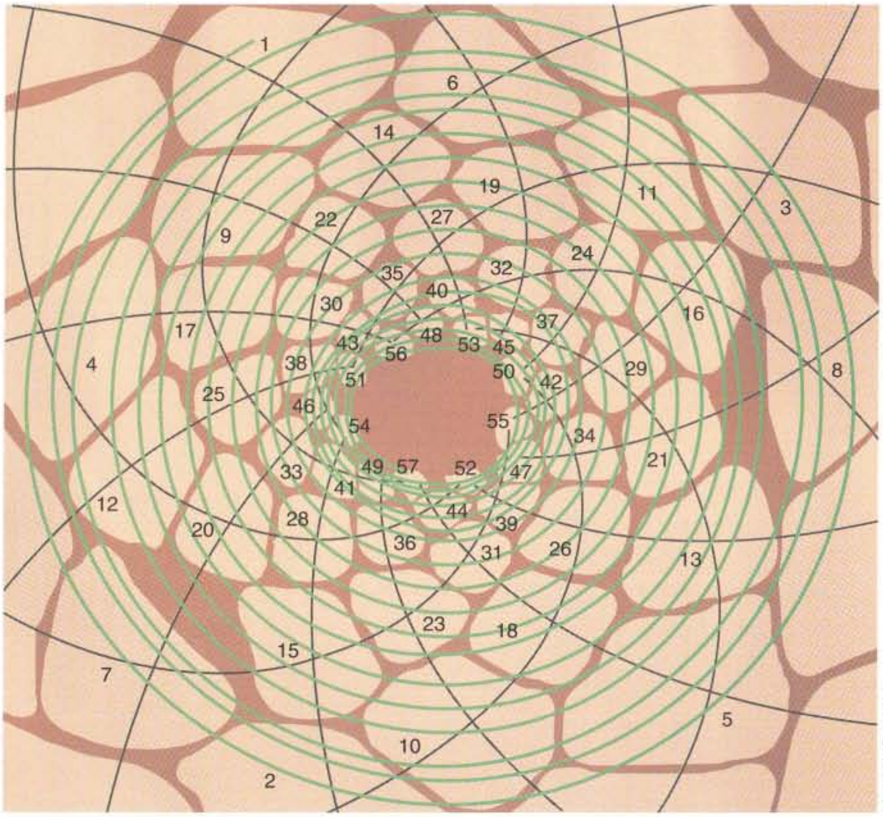
"Well, suppose you did something stupid and used a divergence angle of 180 degrees, which divides 360 degrees exactly. Then successive primordia would be arranged along two opposite radial lines. In fact, if you use any rational multiple of 360 degrees—an angle that can be expressed as $360p/q$ for whole numbers p and q—you get q radial lines and big gaps between them."

Grimes nodded sagely. "So the seeds don't pack efficiently."

"Precisely. To do so requires a divergence angle that is an irrational multiple of 360 degrees—the more irrational, the more efficient. Number theorists have long known that the most irrational number is the golden number."

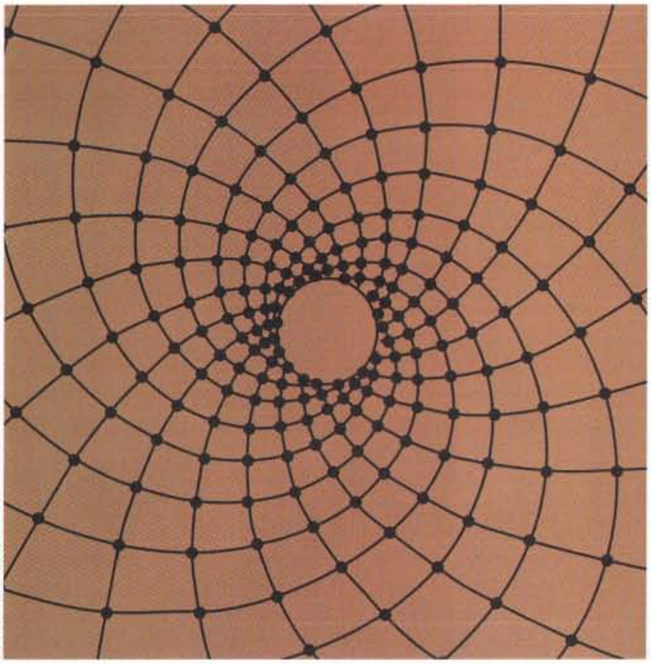
Grimes looked baffled. "What do you mean, 'most irrational'? Numbers are either irrational or not, right?"

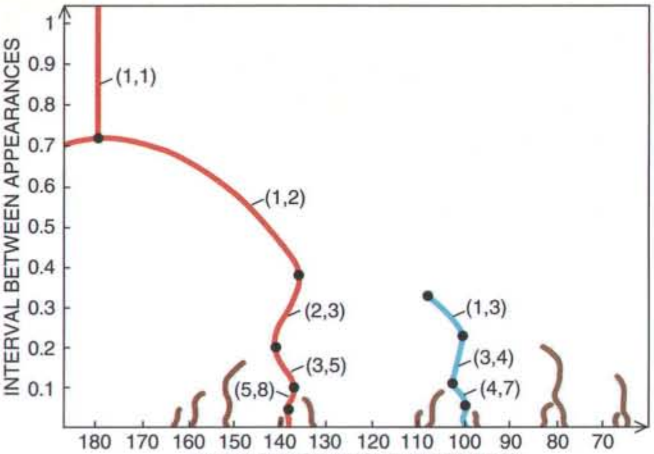
"Yes, but some are more irrational than others. Remember that the ratios of successive Fibonacci numbers tend



JOHNNY JOHNSON

***PRIMORDIA, numbered from 1 to 57, appear in sequence along a tight spiral* (green). *Another 21 spirals, called parastichies, are more obvious. Eight parastichies curve clockwise, and 13 run counterclockwise* (black).**





***ARRANGEMENT of primordia shown here is produced by separating points by the golden angle along a tightly wound spiral* (left). *The number of resultant parastichies depends on the time elapsed between the appearance of successive primordia* (right). *The main curve corresponds to pairs of consecutive Fibonacci numbers* (red). *The secondary curve gives the anomalous series* (blue).**

JOHNNY JOHNSON

toward the golden number φ. So φ is the limit of the sequence $^{2}/_{3}$, $^{3}/_{5}$, $^{5}/_{8}$ and so on. Those are rational approximations that get closer to, but never equal, φ. We can measure how 'irrational' φ is by seeing how quickly the differences between these fractions and φ shrink toward zero. In fact, they shrink more slowly for φ than they do for any other irrational number."

"So the golden number is distinct from every other number based on this simple mathematical property."

"Good point," Bumps said. "Well, Douady and Couder explained the golden angle as a consequence of dynamics, rather than postulate it directly on the grounds of efficient packing. They assumed that successive elements, representing primordia, form at equal intervals of time on the rim of a small circle, or the apex. These elements then migrate radially at some initial velocity and repel one another—a condition that ensures continuous motion and that each new element appears as far as possible from its predecessors."

"You mean it pops up in the biggest gap?" Grimes clarified.

"Grimes, you have a wonderful way with the English language. It's a good bet that such a system will pack efficiently, and so you would expect the golden angle to arise of its own accord. And it does—though with some interesting frills."

"Such as?"

"There are two ways to work out what happens. One is to perform an experiment, as did Douady and Couder. Instead of using plants, though, they filled a circular dish with silicone oil and placed it in a vertical magnetic field. Next they periodically dropped small amounts of magnetic fluid into the center of the dish. The magnetic field polarized the drops, which then repelled one another. To send the drops in a radial direction, they increased the magnetic field at the edge of the dish. The patterns that appeared depended on how much time passed between successive drops. But very often the drops lay on a spiral separated by divergence angles very close to 137.5 degrees."

"The golden angle!" Grimes exclaimed suddenly.

"Douady and Couder produced similar results from computer calculations. In detail, they found that the divergence angle depends on the time elapsed between drops, according to a complicated branching pattern of wiggly curves [*see right illustration above*]. Each section of a curve between adjacent wiggles corresponds to a particular pair of numbers of spiral families. The main branch runs close to a divergence angle of 137.5 degrees, and along it you find all possible pairs of consecutive Fibonacci numbers, in numerical sequence. The gaps between branches represent 'bifurcations' where the dynamics undergo significant changes."

Grimes thought for a few moments. "But there are branches that aren't close to 137.5 degrees, too."

"Yes. The main one corresponds to the anomalous series. The appropriate timing in this same model produces the most common exceptions to the Fibonacci rule and the Fibonacci rule itself—which shows why the exceptions occur while making it clear that they aren't really exceptions at all. But of course nobody is suggesting that botany is quite as perfect as this model. In many plants the rate of the appearance of primordia can speed up or slow down. In fact, whether a primordium becomes a leaf or a petal often accompanies such variations."

"So maybe a plant's genes affect the timing of the appearance of the primordia?" Grimes asked.

"Exactly. But genes need not tell them how to space the primordia. That's done by dynamics. It's a partnership of physics and genetics."

Grimes waved a particularly well-developed sunflower in the air. "You think maybe if I ate these seeds they'd improve my mathematical ability?"

"Try it," Bumps said. Grimes started to nibble at the seeds. "On the other hand—"

He stopped. "What?"

"Don't forget why Fibonacci invented his numbers. Maybe you'll turn into a rabbit."

FURTHER READING

THE ALGORITHMIC BEAUTY OF PLANTS. Przemyslaw Prusinkiewicz and Aristid Lindenmayer. Springer-Verlag, 1990.

PHYLLOTAXIS AS A SELF-ORGANIZED GROWTH PROCESS. Stéphane Douady and Yves Couder in *Growth Patterns in Physical Sciences and Biology*. Edited by Juan M. Garcia-Ruiz et al. Plenum Press, 1993.

LA PHYSIQUE DES SPIRALES VÉGÉTALES. Stéphane Douady and Yves Couder in *La Recherche*, Vol. 24, No. 250, pages 26–35; 1993.

BOOK REVIEWS by Philip Morrison

Smart Buildings

HOW BUILDINGS LEARN: WHAT HAPPENS AFTER THEY'RE BUILT, by Stewart Brand. Illustrated. Viking, 1994 ($30).

By our wastes shall we be known. The most voluminous portion of landfill waste deposits is paper; construction debris is second. Paper dominates since in use it is so ubiquitous yet short-lived. Debris is copious because we build in great bulk and redo what we build. Nationally, the annual cost of all rehab often exceeds that of all new construction (although it earns a smaller share of architects' fees).

This lively, well-prepared critic is penetratingly original, if sometimes carried away by inferential zeal (unlike most others, he admits it). He shows plainly how buildings flow in time. He offers something prospective, too—how and why to build for change sure to come. The hundreds of photographs that crowd this book document it firmly: architecture's frozen music thaws, be it pop, jazz, stately minuet or sacred score. The book presents a reasoned, candid and severe critique of today's architecture and its context in finance and development. At heart this is a refutation of visionary solutions; evolutionary design endures by adapting.

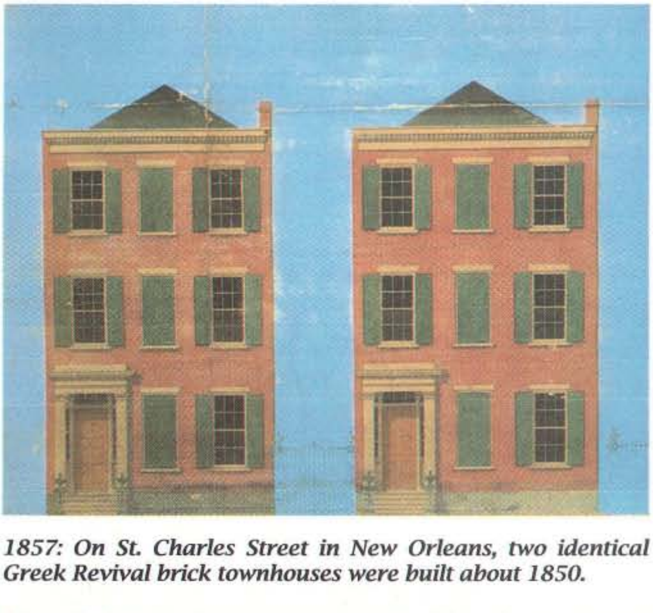
In 1865 New York City's lower Broadway, west side, had 261 buildings from below Trinity Church up to Union Square. Mostly multistoried, they are all drawn here. By 1990 just 33 of them were left, a half-life of 42 years. Cities devour buildings; streets outlive more than 10-fold the structures that line them. In developed nations the largest capital asset, wide open to turnover, is the office buildings, where more than half the workforce is found. Institutional buildings—also mainly offices—change as well, although their businesses rarely fail. Homes change even more systematically as families grow and fade, and desires and opportunities appear. All are pushed by technology, money and fashion.

Every building is a system with multiple rates of internal change. "Site is eternal," says the author, in excusable hyperbole. Structure, skin, services, space plan and the stuff space holds all undergo change at different rates. An adaptive building has to allow for that profound internal slippage, or else the lasting parts will stultify the swift, and quick change will all too soon tear out the durable.

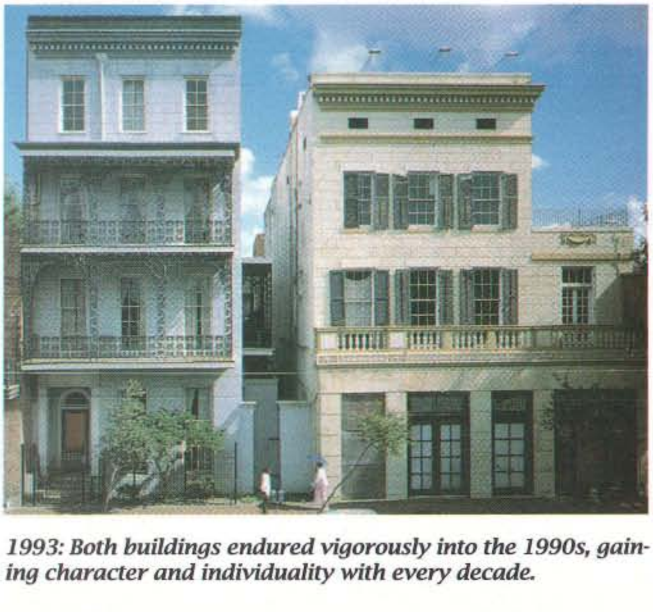
Brand maps three roads to adaptation. The Low Road is freedom: "low-visibility, low-rent, no-style." Nobody cares when you cut openings in sheetrock walls. "Most of the world's work is done" under such a license to change. The founding garages of the start-up firms of Silicon Valley are not mere myths. Even the Pentagon, like the beloved "temporaries" at M.I.T. built during World War II under General Leslie Groves, shares the Low Road by its unpretentious and well-maintained design innocence. The High Road is grander—Monticello, the Devonshires' great house Chatsworth, Robinson Jeffers's Tor House—and can yield subtle adaptations that rest on a continuity of care. "Poets and duchesses know that doing a High Road building right is a labor of love measured in lifetimes."

Lacking both the High virtues and the Low, it is famous buildings that adapt worst. They ignore time, although time does not reciprocate. Here Brand grows scornful, naming renowned names who give us impressive and useless atriums and other alienating designs. The drenched owner of the celebrated Fallingwater designed by Frank Lloyd Wright called it "a seven-bucket building." Most interesting, though, is the author's own disarming mea culpa. He was himself "a major propagandist for Fuller domes" when once they grew in all the wastelands under a young, fond and inexperienced ownership. They always leak, they sound awful, their construction is incurably nightmarish and wasteful, and they cannot adapt at all. Once outgrown, they were mainly left as empty as broken eggshells by their scattered hatchlings.

The architectural profession's studied indifference to adaptation and the real-estate market's indifference to continuity have together induced a "quiet, populist, conservative, victorious revolution," now under way in Britain and America: historic building preservation



1857: On St. Charles Street in New Orleans, two identical Greek Revival brick townhouses were built about 1850.



1993: Both buildings endured vigorously into the 1990s, gaining character and individuality with every decade.

is today an important fraction of all construction. Join this account to a chapter on maintenance, and the text has entered the design domain. One aphorism is well worth citing: "The Eiffel Tower's lasting message is: exposed structure can be gorgeous." The inside-out design of Pompidou Centre has a lasting message as well: "Never expose services." All those brightly painted ducts, funnels and pipes are less than rigid under wind and weather. Once slightly bruised or bent, they soon proclaim an "exorbitant scandal of rust and peeling paint."

Pride of place in a chapter on vernacular styles is given to the Malay House, "a wonder of incremental architecture." Around it there has developed a specific language of growth: start with the core house, go to verandas, a covered walkway, kitchen, extensions front and back, and entrance porch. In the monograph cited here ("the best book that I've seen on any indigenous architecture"), its author Lim Jee Yuan credits the style with "near-perfect solutions to the control of climate, multifunctional use of space, flexibility in design and a sophisticated prefabricated system which can extend the house with the growing needs of the family." One such house is shown here, beautiful inside and out; you see a tight steel roof shielding an open, airy tolerant house on stilts, with a sense of generous order.

Other chapters continue in the same winning voice, an iconoclasm that is never merely contentious. End with one sidelong passage on materials. "I remember when fiberglass boats first came along...and everyone said they would never work, never sell, never last. Wrong on every count. Fiberglass boats are lighter and stronger...[and] endure negligent owners.... Fiberglass never leaks, wood always does.... I have owned and sold an excellent plastic boat and owned and kept a troublesome wooden boat. Why? The wood feels better, and I can fiddle with it. But if I really had to sail somewhere I'd get fiberglass or steel."

Buildings will follow that course, Brand believes. "The ones that have to sail somewhere will be made of advanced materials." Traditional materials will keep their costly appeal, for those who will pay for a touchable and rewarding aesthetic. All around us will grow a clever second domain of the more practical, mostly proud and absolute fakery, like granite veneer. But future buildings will learn a little more easily, by paying attention as we do, their new smart materials provided in one way or another with built-in sensory and motor responses. This is a readable book on learning, itself good to learn from—and to dispute.

Autofab

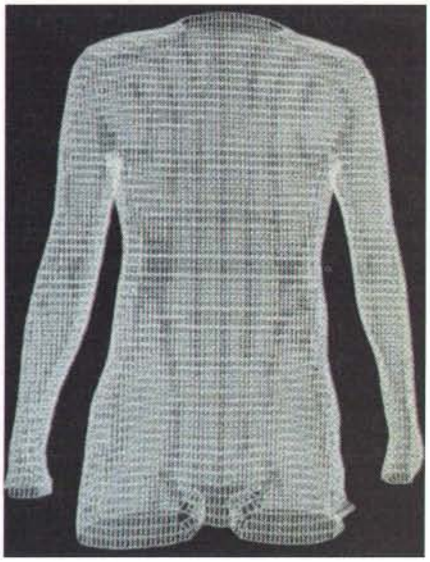
AUTOMATED FABRICATION: IMPROVING PRODUCTIVITY IN MANUFACTURING, by Marshall Burns. PTR Prentice-Hall, 1993 ($64).

This fresh, readable, often surprising book looks at autofab, the manufacture of complex shapes without assembly of parts, with minimal human attention and without a new solid model for each new shape. Autofab can do without solid prototypes; its models are digital.

It is easy to be a sculptor, runs the old challenge. Simply take away every bit of your marble block that is *not* Venus. That subtractive shaping of solids is still the chief task of machine tools. *Subtractive* autofab is not new. It began about 40 years ago; by the late 1980s it was widespread, mainly in the form of computer numerically controlled (CNC) milling, an engineering industry now approaching the $10-billion-a-year scale. In such machines, rotating tools cut away the workpiece under tight digital command. One giant instance is shown here, a 1992 U.S.-made CNC "machining center," ready to set its axis "over any machinable solid," to an accuracy of five or 10 microns, anywhere within a volume the size of a large living room. At any point it will select one tool from its kit of 100 cutters, drills and taps, pausing only for "60 sec chip-to-chip." For $5 or $10 million, you can own one fit to release the frame for some heavy earthmover, or the Pallas Athena herself, from a slab of metal or stone.

Half a dozen other subtractive forms of autofab, none so heroic in scale, are described here quite fully. One firm makes a machine able to carry out unattended point-by-point shaping entirely by CNC electrical spark erosion and thus to form with precision even substances that resist cutting tools, such as hardened steels and certain ceramics. The most modest CNC device listed is a trainer for student machinists, a desktop lathe under full PC control, priced at 1.6 kilobucks by its Austrian makers.

That machine-tool style of autofab is vaguely familiar to most readers. But it is the new *additive* autofab that has riveted the author's attention and in turn wins ours. Additive autofab is an embryonic industry, now in boom, on an economic scale still only mills to the dollar of its mature subtractive rival. There Sony, DEC, Du Pont and others strive among start-ups large and small.



SURFACE RELIEF of human body digitized by a planar interferometer can be used in the automated fabrication of garments, especially those for outer space and other harsh environments.

Among several less ambitious systems now in use, one product here is signaled by a working demo "generated automatically in a single machine" in 24 hours of work. This desktop showpiece is an assemblage of 12 white plastic gears, so meshed that turning any one rotates all the others. A plastic universal joint offers a more practical example. Using an automatic copier-like system of high quality, the system begins by printing out on glass, full-size pattern masks the final product design, sectioned into layers. Fluid resin layers are laid down at about one per minute, each layer in turn cured into solid by exposure to ultraviolet light. The uncured portions left in the shadows are sucked away. Molten wax fills the voids to support the delicate structures in the layer. The wax is chilled and milled flat before the next layer is placed down, to improve accuracy—kept within some five or 10 mils—and to strengthen the product. Finally, that wax "scaffolding" is removed, a task still not fully automated.

Several variants are in use. The most popular cures the polymer by laser scanning point by point, although potentially exposing an entire layer at a time is much faster. The resin used is critical, for it constitutes the final product, but photosensitive polymers hardly span a wide range of materials. So far most users produce prototypes meant for study and test, for molds, 3-D maps, models, even art. One new additive scheme emplaces layers of powder onto which droplets of a binder fluid are artfully deposited very much as letters are

SCIENTIFIC AMERICAN

COMING IN THE FEBRUARY ISSUE...

TRENDS IN SEMICONDUCTOR MANUFACTURING

Gary Stix,
staff writer

THE HISTORY OF TESTOSTERONE

John Hoberman
University of Texas at Austin

Charles Yesalis
Pennsylvania State University

CREATIVITY AND MADNESS

Kay Redfield Jamison
Johns Hopkins University School of Medicine

ALSO IN FEBRUARY...

Population, Poverty and the Environment

Masers in the Sky

Molecular Machinery That Drives Gene Regulation

Sonoluminescence

The Mid-Cretaceous Superplume Episode

ON SALE IN FEBRUARY

formed on paper by ink-jet printers. Using alumina powder and silica adhesive, these forms are suited for the kiln, making ceramic mold shells for precision investment casting with no need for any initial pattern. The commercial version is now in tryout. Five or six other processes are in production, of which selective sintering by laser scan of deposited layers of powder sounds the most likely.

The author, a physicist turned computer entrepreneur, writes well. He has managed the rare feat of giving plenty of timely commercial detail without becoming boring, by use of clear tables, visual examples and lively comment. More technical chapters treat choice among producers, data-input methods and a range of materials.

With some help from his optimistic friends, he offers philosophical essays on the social and economic impact. "Customer coproduction" of automobile body design by computer screen choice is easily forecast, although as a revolution it does seem effete. Given millionfold users of Personal Autofab, PA-savvy hosts could welcome guests with special dinner plates or trendy trinkets made of recyclable powders from that evening's newly downloaded software. It is not easy to extrapolate to a time when automatic fabrication is as commonplace as SimCity; can any PA hope to lay down zirconia paring knives? The layered future lies still in shadow, perhaps awaiting chemical systems or even atomic beams.

Genetic Distance

THE HISTORY AND GEOGRAPHY OF HUMAN GENES, by L. Luca Cavalli-Sforza, Paolo Menozzi and Alberto Piazza. Princeton University Press, 1994 ($150).

Our genes stipulate just how biological Nature guides societal Nurture. Yet the fascination of this overflowing treasury of historical genetics is its unexpected demonstration that Nurture, obedient not to inborn nucleotides but to the intricacies of culture, moves along with Nature on one shared path. When two travelers walk hand in hand, who is guide?

This outsize volume is a hardcover match to the bulk of Boston's *Yellow Pages.* It reviews at depth 14 years of pathbreaking teamwork by this trio of scholars from Palo Alto, Parma and Turin. The authors' data are drawn from a wide literature, thousands of reports over four decades of tests, mostly immunologic results on blood samples, taken in 7,000 locations among 1,900 named peoples, Ainu to Zuni. An average sample might test for 10 genetic traits among a few hundred individuals drawn from unrelated local families. All groups sampled are those we call aboriginal, living today more or less where they were at the time of Columbus. They, too, have usually wandered but in migrations that predate by millennia the swift dispersal of Europeans after the 15th century.

The trick is to judge genetic kinship by easy, opportunistic tests. Gross visible physical features such as height and weight are under strong extragenetic influence. Eye, hair and skin color are genetically complex and have proved of little value for analysis. But the proteins on the surfaces of blood cells are easy to characterize tightly by the response of small samples to specific antibodies. The familiar blood groups that govern transfusion derive from red cells. A much larger family of distinct genetic markers—most of the 100 used here—has more recently been identified in the rich antigenic properties of white lymphocytes.

These rather random marker proteins are not in themselves very significant. But they allow quantitative study of inheritance within any group. (Recent mutations are rare and add only a little noise.) The dozen gene forms, or alleles, of the most important single marker used, a white cell surface protein called HLAA, are mapped and tabulated as they were found worldwide, and so on for almost 100 other alleles.

Some alleles are all but universal. They tell us little beyond their selected value to our common humanity. Some are rare save in the small region where they arose long ago. But most are found everywhere at some level of frequency. Thus, single alleles can say little. They are apparent in nearly all populations. On that rock the old and wicked idea of human races foundered utterly. People vary genetically within groups more than they vary from one group to another; most evolution took place during the long time before humans made it to all the continents.

The key tool is a quantitative means to measure differences in the *pattern* of gene frequencies, not in any single one. What is sought is a measure of genetic distance between two groups that weighs many gene frequencies in a statistically unbiased way. The mathematics, treated here at length, is neither simple nor unique. But the outcome is persuasively tested in several different ways. One reassuring result is the remarkably smooth relation obeyed by the "genetic distance" between pairs of groups and the simple geographical distance between the group locations.

A second test is applied to the formation of "trees of descent" among groups, again taking genetic distance as a measure. The two groups of least distance form the lowest branch of the whole tree. That pair is pooled, and new distances found to all others. Finally, only two populations are left; they define the earliest fission.

The authors discuss in turn continent after continent. The bulk of the text is given to these detailed chapters. The many genealogies confirm one another. The most robust tree arrays 42 major groups pooled from the full database. About 120 alleles are taken into account. The first split neatly parts Africa from all the rest. On the non-African branch the next split divides north from south Asia. The Amerind groups then fission from Northeast Asia, and that Asian branch itself ends in a split between the Arctic peoples and the Tibetan, Japanese, Mongolian and Korean peoples.

The linguists, too, have a huge database, words instead of genes, in 6,000 tongues. Their latest essay at one family tree of human tongues—by no means yet anywhere near consensus—strongly parallels the tree of alleles. A residual subjectivity in graphical form does not lessen the remarkable fit between 16 linguistic families and the 40-odd genetic populations: each language family can be quite well matched to one or to a very few of the closely related genetic clusters.

Try a couple of less ambitious comparisons. On maps of Spain, many local place-names and some Neolithic cave paintings, stone tools and cranial remains all conform to clear gene patterns: the influence of the Basques is marked in all, spreading over most of Iberia with a strength that declines as distance grows from the present Basque country.

Consider three sociological types of French families: nuclear, patriarchal and extended, all of them identified more than a century ago. The three types are still geographically distinct. Their distribution over France correlates well both with today's voting patterns and with gene frequencies. Blood group O and the socialist vote in France are closely correlated in location, because the family structure, like the biological heritage, is reliably carried by young children, and adult opinion follows. It is not that red cell proteins mysteriously fix voting habits but rather that genes and cultures walk hand in hand. Here is a comprehensive account of how well we can follow them both into the long, still indistinct past.

In the World's Developing Countries, we're planting trees for less than…

…*A NICKEL EACH!*

If you'd like to help, please contact:

TREES for the FUTURE
11306 ESTONA DRIVE. P.O. BOX 1786
SILVER SPRING, MARYLAND 20915-1786

1(800) 643-0001

AUDIO-FORUM® offers the best in self-instructional foreign-language courses using audio cassettes — featuring those used to train U.S. State Dept. personnel in Spanish, French, German, Arabic, Russian, Hebrew, Chinese, Italian, Korean, Polish, and more.

Learn a foreign language on your own!

Free Catalog.

Our 23rd year.

Audio-Forum,
Room E139, 96 Broad St.,
Guilford, CT 06437 (203) 453-9794
Fax: (203) 453-9774
e-mail: 74537.550@Compuserve.Com

Name

Address

City/State/Zip

I am particularly interested in (check choices):
☐ Spanish ☐ French ☐ German ☐ Russian
☐ Hebrew ☐ Chinese ☐ Arabic ☐ Polish
☐ Other

SCIENTIFIC AMERICAN

Cumulative Index on Computer Disk

Available for Macintosh, and IBM and compatibles, running under Windows or DOS.

Only $49.95

Order SciDex® today and turn your *Scientific American* library into an invaluable reference tool. Includes full documentation.

- ***Article Abstracts***
- ***530 Issues***
- ***4,300 Articles***
- ***Over 5,000 Authors***
- ***Print Your Search***
- ***43,000 Topic Entries***

GUARANTEE: If your copy of SciDex is defective, return it with your registration number within 90 days and we will promptly replace it free of charge.

SCIENTIFIC AMERICAN SciDex®

415 Madison Avenue
New York, NY
10017-1111

Please send me _____ copies of SciDex, the *Scientific American* electronic index from May 1948 to June 1992 at $49.95. Add $5.00 for domestic shipping and handling.* Corporate orders accepted if accompanied by authorized purchase order. Allow 4 to 6 weeks for delivery. Be sure to select version and disk format below.

Name ____________
Organization ____________
Address ____________
City ____________ State ______
Zip ______ Fax: (   ) ______
Tel: (   ) ____________ Please Ship:
☐ Macintosh® ☐ Windows™ ☐ MS-DOS® version as
☐ 3-1/2″ DS/HD,
☐ 3-1/2″ DS/DD (Not available for Windows™)
☐ My check/money order is enclosed for $______
Charge my ☐ VISA ☐ MasterCard
☐ ACCESS ☐ EuroCard Exp. Date ______
Card No. ____________
Signature ____________

*Add applicable sales tax for IA, IL, MA, MI, CA, NY, PA. In Canada, add $3.50 for GST (No. R127387652). SD13

Outside the U.S. remit $49.95 in U.S. funds drawn on a U.S. bank or by credit card and add $5.00 for surface delivery and handling or check here ☐ and add $15 for air delivery and handling. International and Domestic Credit Card orders accepted by fax (212) 980-8175.

SYSTEM REQUIREMENTS: SciDex can be used on any Macintosh (Mac Plus or better) with 2Mb RAM (4Mb recommended), running under System v6.0.5 or later and a hard disk with 5Mb of free space. SciDex comes compressed on high-density 1.4Mb or double-density 800K 3-1/2″ disks (800K not available for Windows). The Windows and DOS versions require an IBM compatible computer. For the Windows version: Microsoft Windows 3.0 or later; one megabyte of memory (an 80386 with 2Mb is recommended); a hard disk drive with 8.5Mb of free space; and an EGA or higher resolution monitor. The DOS version requires MS-DOS 3.0 or later; 256K of RAM; a hard disk drive with 4.5Mb of free space; and a monochrome or higher resolution monitor. Please specify disk format.

ESSAY by John Timpane

How To Convince a Reluctant Scientist

> At the end of reasons comes persuasion.
> —Ludwig Wittgenstein, *On Certainty*

At the end of every successful argument, no matter how weighty the evidence or powerful the reasoning, the beholder performs a nonrational act: the leap to acceptance. It may be short—one may feel pushed—but a leap it always is.

What makes a scientific argument persuasive? Granted, there is no one nation called "science" but a panoply of tribes with their respective handicrafts. A cosmologist might nod, and a chemical engineer squirm, when someone argues that the unobserved *must* exist for the observed to make sense. Certain branches of physics absolutely demand reproducibility—but an anthropologist may get only one chance to live amid the !Kung.

There are some basic persuaders, more or less familiar, more or less honored. Robert K. Merton's five elements of scientific endeavor—originality, detachment, universality, skepticism and public accessibility—all play a role. If results and methods follow the rules we know, and if we find the candor and testability we expect, we are at least nearer to feeling well disposed. Yet in the sciences as elsewhere, conviction involves other forces, both reason and nonreason.

Our sense of what connects, what fits, what ramifies, can often move us to acceptance. Results that bear little connection to existing work will persuade little. But if study A forges connections between studies B, K and S, or if it newly illuminates an established fact, a tiny epiphany can happen.

Or a massive one. Sir Isaac Newton's audience was impressed when he derived his laws of physics from observations made by Galileo—but overwhelmed when he then showed that the moon, planets and stars obeyed these same principles. It did not escape the notice of James D. Watson and Francis Crick that their proposed structure for DNA helped to explain the copying mechanism of the genetic material. Ramification may work also from without. If we realize the argument we have just heard is supported by independent evidence, our personal paradigms may start to budge. Some scientists initially resisted John Ostrom's argument that the ancestors of modern birds were ground-dwelling theropods. But when we watch a hawk hop after a desert rat, we are not far from the awkward leap of *Archaeopteryx*.

To be persuasive, then, you can be only so original. Total originality, to the point at which no one else speaks your language, may cripple your argument's progress in the world (as happened initially to Thomas Young's wave theory of light). Call it the Law of Obviousness: better to be a half-step ahead and understood than a whole step and ignored.

Prestige sells, too. Freeman Dyson has noted the "profound consequences" of status in the sciences. A famous team at a famous institution backed by famous money is a hard combination to resist. Even harder is the status that comes with large projects. Almost irresistible is the assent of prestigious colleagues or superiors.

Publicity, for better or worse, begets prestige. Scientists today—of necessity—are veritable master organists at the great keyboards of the media. Going public spurs debate, encourages verification, attracts funds. Today, when there is more science than ever before, a number of scientists see no alternative to invoking the hosanna and the gee whiz.

Many scientists believe arguments they have never read or heard, simply because most of their peers believe. As Thomas S. Kuhn writes, "There is no standard higher than the assent of the relevant community. The transfer of allegiance from one paradigm to another is a conversion experience that cannot be forced."

And sometimes the sheer beauty of a demonstration can render an argument compelling. Einstein's special theory of relativity gained acceptance not only by its power to explain observed phenomena but also by the elegance of the equation $E = mc^2$. Cell biologists wax rhapsodic over Erkki Ruoslahti's experiments showing that a certain molecule, called an integrin, mediates the adhesion of cells to the protein fibronectin. Ruoslahti's work was of such beauty that other cell biologists found it extremely and immediately persuasive.

Much of what makes a scientific argument convincing, in the end, has to do with things other than science. By saying so, we do not impugn or undermine the sciences so much as we recognize their humanity. Some postmodernist thinkers—those, for example, who practice the collection of methods called deconstruction—believe that the important role of the nonrational in the sciences is a scandal, when in fact that role is fairly familiar to anyone who actually does a science for a living.

Besides, such conceptual issues cease to have much importance past the lab door, which opens into a world of different priorities. In the workplace, scientists have utter faith in the reality of the world and in themselves. They have to. (Otherwise, their experiments, and perhaps they, may turn to tar.) The order of things, furthermore, pressures them to treat all data, especially their own, with thorough skepticism. After all, these pragmatic, naive, realist skeptics must eventually submit their findings to another entity in which they trust—the candid, bruising machinery of peer review.

How, then, does the history of science unfold? According to reason, drawn by the ineluctable, magnetic pull of truth? Or by bursts, gaps, sudden nonrational disjunctions? Perhaps the latter amounts to the former, since the aggregate of thousands of nonrational moments—when individual scientists are persuaded, convinced, converted—adds up to a progress of sorts. Conversion happens when a piece falls into place and renders the whole puzzle new. Often the new vision is so powerful that our decision to accept may seem hardly a decision at all. Yet a decision it always is. Precisely because they are not rational, such leaps—from final ice floe to riverbank—are wonderfully, deeply human.

JOHN TIMPANE is a writer who teaches English at Lafayette College in Easton, Pennsylvania.

150 YEARS OF INNOVATION **SCIENTIFIC AMERICAN**

Special Issue. Order Now.

Sponsored by Apple

"The Computer In The 21st Century"

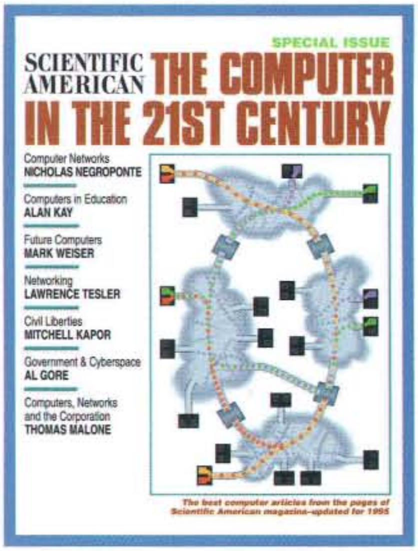
If you're a traveler along the world's information superhighway – or anticipate venturing down that road – be sure to use this trail-blazing guide from *Scientific American*: "The Computer In The 21st Century."

This unique anthology features 32 acclaimed *Scientific American* articles by the world's leading experts, including:

- Computers, Networks And Work by *Lee Sproull, Boston University* and *Sara Kiesler, Carnegie Mellon University*
- Networked Computing in the 1990s by *Lawrence G. Tesler, Apple Computer Inc.*
- The Computer For The 21st Century by *Mark Weiser, Xerox Palo Alto Research Center*
- Lightwave Communications by *Emmanuel Desurvire, Columbia University*

"The Computer In The 21st Century" features over 125 pages of information on lightwave communications... microlasers... cyberspace... virtual reality... quantum cryptography... digital technology... and so much more. All beautifully illustrated with the colorful art and graphics for which *Scientific American* is renowned.

Don't let the future pass you by – request your copy of "The Computer In The 21st Century" by returning the coupon today with your remittance. Consider ordering several issues – they make great gifts for friends and colleagues.



SPECIAL ISSUE

SCIENTIFIC AMERICAN THE COMPUTER IN THE 21ST CENTURY

Computer Networks
NICHOLAS NEGROPONTE

Computers in Education
ALAN KAY

Future Computers
MARK WEISER

Networking
LAWRENCE TESLER

Civil Liberties
MITCHELL KAPOR

Government & Cyberspace
AL GORE

Computers, Networks and the Corporation
THOMAS MALONE

The best computer articles from the pages of Scientific American magazine–updated for 1995

AVAILABLE AT NEWSSTANDS IN MARCH

Avoid Disappointment – Reserve Your Copy Now!

Mail to: **SCIENTIFIC AMERICAN** Dept. CA, P.O. Box 11314
Des Moines, IA 50340-1314

Please reserve my edition(s) of "The Computer In The 21st Century." My check or money order, payable to *Scientific American*, is enclosed.

Name (PLEASE PRINT) ____________________

Company ____________________

Address ____________________ Apt./Suite ________

Post Code ________ City ____________ Country ________

Please send _______ copy(ies) at $3.95 each $________

Plus $1 per copy postage & handling (1-9 copies) $________

TOTAL AMOUNT ENCLOSED $________

Please remit in U.S. funds drawn on a U.S. bank.

PBCAF5A

Reach the innovators of Japan through the pages of Nikkei Science, Scientific American's Japanese-language edition.

With a circulation of 27,096, Nikkei Science's pages host a unique audience:

Median Age	35.7
Household Income $48,000+	60%
Graduated from 4 year college or more	80%
Median time spent reading NIKKEI SCIENCE	3.6 hours

Types of computers used by subscribers:

Personal	77.2%
Workstation	26.6%
General-purpose	21.1%
Other	27.0%

For further advertising information contact:

Mr. Tsutomu Nakajima
NIKKEI SCIENCE
No. 2-1, 2 Chome
Uchisaiwaicho, Chiyoda-ku
Tokyo, Japan
Telephone: 011-81-3-5512-1546
Fax: 011-81-3-5511-8578

THE JAPANESE-LANGUAGE EDITION OF SCIENTIFIC AMERICAN

NIKKEI SCIENCE